Trends In Amplification

VOLUME 8, NUMBER 2, 2004

Hearing Aids and Music

Marshall Chasin, AuD, and Frank A. Russo, PhD*

Historically, the primary concern for hearing aid design and fitting is optimization for speech inputs. However, increasingly other types of inputs are being investigated and this is certainly the case for music. Whether the hearing aid wearer is a musician or merely someone who likes to listen to music, the electronic and electro-acoustic parameters described can be optimized for music as well as for speech. That is, a hearing aid optimally set for music can be optimally set for speech, even though the converse is not necessarily true. Similarities and differences between speech and music as inputs to a hearing aid are described. Many of these lead to the specification of a set of optimal electro-acoustic characteristics. Parameters such as the peak input-limiting level, compression issues—both compression ratio and knee-points—and number of channels all can deleteriously affect music perception through hearing aids. In other cases, it is not clear how to set other parameters such as noise reduction and feedback control mechanisms. Regardless of the existence of a "music program," unless the various electro-acoustic parameters are available in a hearing aid, music fidelity will almost always be less than optimal. There are many unanswered questions and hypotheses in this area. Future research by engineers, researchers, clinicians, and musicians will aid in the clarification of these questions and their ultimate solutions.

1. Introduction

Music as an input to a hearing aid is a relatively new concept for many hearing health care professionals and hearing aid design engineers. Although technical innovation and, in many cases, ingenuity have resulted in some very useful devices for the hard of hearing, most are predicated on speech as input rather than music. For some types of music, the speech-input settings yield electroacoustic characteristics that are near optimal. However, these settings can be less than optimal for most other types of music.

Not only is the technology for a music-input still in its infancy, but the research and clinical knowledge of what musicians and those who like to listen to music need to hear is also still in the early stages of understanding. Nevertheless, some strategies have been clinically shown to be useful for many musician groups. In some cases, there

From the Musicians' Clinics of Canada, Toronto, Ontario and *University of Toronto, Mississauga, Ontario, Canada.

Correspondence: Marshall Chasin, AuD, Audiologist, Musicians' Clinics of Canada, 340 College Street, Toronto, Ontario, M5T 3A9; E-mail: Marshall.Chasin@rogers.com.

©2004 Westminster Publications, Inc., 708 Glen Cove Avenue, Glen Head, NY 11545, U.S.A.

may be clear statements of the need for more research and in others, an abrogation of clinical responsibility to the musician: "try this out and tell me what changes you might like."

Unfortunately, hearing health care professionals and musicians use a different terminology. For example, an audiologist may say "440 Hz" and a musician may say "A" (or in Canada, "eh"). This has long been a stumbling block between musicians and hearing health care professionals, but is not insurmountable. Musicians are highly interested in notes on the left side of the piano keyboard below middle C (262 Hz). Most clinicians would rather ignore the sound energy below 250 Hz because of the poor signal-to-noise ratio and because of hearing assessment problems. However, for some musicians, this low-end information can significantly contribute to the quality of music.

A cornerstone of any clinical practice is to have a joint exploration of the hard-of-hearing musician's needs using both audiologic and musical knowledge. It is not necessary for the hearing health care professional to have an in-depth knowledge of music and the converse is also true. A clinical partnership between audiologist and musician tends to yield the best results.

The following discussion concerns the special issues involved in hearing health care where music is a concern. Topics include speech and music as input, the parameters of music, hearing aids, and other hearing assistive devices. The issues that have been identified and the approach to hearing health care discussed may be relevant for hard-of-hearing musicians as well as for nonmusicians who like to listen to music. This equally applies to rock, blues, jazz, grunge, Mozart and Beethoven.

2. Differences Between Speech and Music for Hearing Aids

Just as there are similarities and differences between speech and music spectra, similarities and differences also exist between the perceptual requirements for speech and music. Compared with music, speech tends to be a well-controlled spectrum with well-established and predictable perceptual characteristics. In contrast, music spectra are highly variable and the perceptual requirements can vary based on the musician, type of music, and the instrument being played. At least five salient differences between speech and music have direct ramifications for hearing aid fittings.

2.1. Speech Versus Music Spectra

Speech, regardless of language, has to be generated by a rather uniform set of tubes and cavities. The human vocal tract is approximately 17 cm from larynx (vocal chords) to lips (Kent and Read, 2002). The vocal tract can be either a single tube, as is the case of oral consonants and vowels, or a pair of parallel tubes when the nasal cavity is open as in [m] and [n]. Nevertheless, fundamental laws of acoustics that are independent of spoken language govern the human vocal tract. All sound that is emanated from the vocal tract is affected by the damping characteristics of the cheeks, tongue, and the nasal cavity; resonant characteristics that are both wavelength and Helmholtz related; and constrictions by complex but well-defined articulators.

For example, formants (the resonant frequencies of the vocal tract) are governed primarily by constrictions in the mouth and the length of the vocal tract tube. Vocal tract lengths cannot change significantly; thus, it is understandable that the adult vocal tract generates a rather limited set of outputs. Taken together and measured over a period of time, this can be summarized as the long-term speech spectrum. Hearing aid engineers and hearing health care professionals have sought to reestablish the shape of this spectrum for hearing-impaired listeners, via amplification, with hopes of improving speech communication. Indeed, many of the target-based hearing aid-fitting formulae are based on the long-term speech spectrum.

In contrast to the relatively well-defined human vocal tract output, the long-term music spectrum likely resembles low-pass filtered noise. The long-term music spectrum is a relatively meaningless norm, however. The outputs of various musical instruments are highly variable, ranging from a low-frequency preponderance to a high-frequency emphasis. In some cases, the output spectrum is "speech like" whereas in others, there is no similarity. Essentially, there is no single music-based target that can be the goal of an optimal hearing aid fitting.

2.2. Physical Output Versus Perceptual Requirements of the Listener

In speech, slight differences exist among various languages in the proportion of audible cues that

are important for speech perception. This has been summarized under articulation index (AI) research. Measures such as the AI have been used for decades in the hearing aid industry. The AI weightings as a function of frequency do vary slightly from language to language but generally show that most of the important sounds for speech clarity derive from bands over 1000 Hz, whereas most of the loudness perception of speech is from those bands below 1000 Hz. Clinically, it is accepted that if a client reports unclear or muffled speech, a decrease in low-frequency and/or an increase in high-frequency sound transmission will generally help alleviate the complaint.

One may say that with speech, the spectrally most intense region with most of the energy is in the lower frequencies; clarity, which has more to do with consonants, is derived from the higher frequencies. Linguistically speaking, speech is phonetically more dominant in the lower frequencies and is phonemically more important in the higher frequencies. That is, the auditory perception of speech has a significantly different weighting than does the physical output from a speaker's mouth. Despite the differences between the physical output of the speech and the frequency requirements for optimal speech understanding, the differences are constant and predictable: low-frequency loudness cues and high-frequency clarity cues. The need to accomplish an appropriate balance between speech loudness and speech clarity has important ramifications for fitting hearing aids.

Unlike speech, the phonemic spectrum of music is highly variable. The perceptual needs of the musician or listener may vary depending on the instrument, regardless of its physical output. A stringed instrument musician needs to be able to hear the exact relationship between the lower-frequency fundamental energy and the higher-frequency harmonic structure. The violinist who says, "this is a great sounding instrument," is often referring to the relationship between the fundamental energy and the harmonics having a preferred balance both in relative intensity and exact spectral location. One can say that a violinist therefore has a broadband phonemic requirement up to 6000 Hz. A violinist not only generates a wide range of frequencies but also needs to be able to hear them.

In contrast, a woodwind player such as a clarinetist needs to be able to hear the lower frequency (*eg,* 1500 Hz) inter-resonant breathiness. A clarinet player who says "that is a good sound" is saying that the lower-frequency sound in between resonances of the instrument has a certain level. Overly sharp resonances are not desirable, and high-frequency information is not very important to a clarinet player other than to assist with loudness perception. One can, therefore, say that a clarinet player has a low-frequency phonemic requirement even though the clarinet player can generate as many higher-frequency sounds as the violinist. Setting a hearing aid to provide each of these musicians with the optimal sound would be different exercises—the violinist needs a speech-like broadband-aided result, whereas the clarinet player would be just as happy with a 1960s hearing aid frequency response. However, fitting both of these musicians with a broadband modern hearing aid would be acceptable, if not superfluous (in the clarinetist case).

2.3. Loudness Summation, Loudness, and Intensity

The source of sound in the human vocal tract is the vibration of the vocal cords. Because of the way the vocal cords are suspended in the larynx, they function as a half-wavelength resonator (*ie,* held rigidly at both ends). This means not only is there energy at the fundamental frequency (typically 120 to 130 Hz for men and 180 to 220 Hz for women), but harmonics are evenly spaced at integer multiples of the fundamental frequency. A man's voice with a fundamental frequency of 125 Hz has harmonics at 250 Hz, 375 Hz, 500 Hz, and so on. Fundamental frequencies are very rarely below 100 Hz; therefore, the minimal spacing between harmonics in speech is on the order of at least 100 Hz. In other words, no two harmonics would fall within the same critical band, with the result that there is maximal loudness summation—a soft sound has a low intensity and a loud sound has a high intensity. With speech, there is a good correlation between one's perception of the loudness and the physical vocal intensity. Setting a hearing aid to reestablish equal or normal loudness with speech is therefore a relatively simple task.

Some musical instruments are speech-like in the sense that they generate mid-frequency fundamental energy with evenly spaced harmonics. Oboes, saxophones, guitars, and violins are in this category. Some, such as the clarinet, are quarter-

wavelength generators (odd numbered harmonics), at least for the lower frequency notes. Many bass stringed instruments, such as the string bass and the cello, are half-wavelength resonator instruments similar to speech but tend to be perceived as less loud since more than one harmonic can fall within one critical bandwidth and thereby not contribute to the loudness because of the lack of loudness summation. Recall that if energy is constrained within a critical band, the addition of other elements within that band does not contribute to the subjective perception of loudness. This is the basis behind loudness summation. All energy components within a critical bandwidth can be thought of as being in one bucket, and the addition of more energy to that one bucket will not alter the loudness. Contribution to more than one adjacent bucket (*ie,* additional energy components in a frequency region that is in excess of the critical bandwidth) will contribute to the perception of increased loudness.

Because the bass and cello produce tones with more than one harmonic in a critical band, the correlation is poor between measured intensity and perceived loudness—minimal increase in loudness as a function of increase in playing intensity. A hard-of-hearing bass or cello player would therefore need less low- and mid-frequency amplification to reestablish equal or normal loudness perceptions than those who play other musical instruments. Instruments with more treble, such as the violin and flute, have each of their harmonics in different critical bandwidths, thus adding to the perception of loudness. A good correlation between measured intensity and perceived loudness exists with these instruments.

A music program for bass and cello players would need to be set with less low- and mid-frequency gain than one for players of treble-oriented instruments. Failing to do so would cause the bass notes of the music to sound overly loud for cello and bass players such that they would not play as intense. Although this might contribute to a lowering of the overall intensity of the orchestra, the bass and cello players would not be able to adequately hear their own instrument and might use nonergonomic head positions, leading to subsequent neck strain. This might include leaning their ears so as to contact their instruments' tuning pegs to improve monitoring.

2.4. The "Crest Factor" of Speech and Music

The crest factor is a measure of the difference in decibels between the peaks in a spectrum and the average or root mean square (RMS) value. A typical crest factor with speech is about 12 dB. That is, the peaks of speech are about 12 dB more intense than the average values. This is well known in the hearing aid industry and was the basis for the reference test gain measure in various versions of the ANSI hearing aid performance standard. Damping is one of the physical parameters that led to the 12-dB crest factor for speech. The human vocal tract is highly damped. The highly damped nasal cavity, soft cheeks, soft tongue, lips, and saliva all contribute to a highly damped (if you'll excuse the pun) output. One of the many reasons for setting the threshold knee (TK)-points on hearing aid compression systems is predicated on this crest factor. Compression detectors are set to function according to this 12-dB crest factor.

Musical instruments are not so well damped, however. Hard walled horns and stiff resonator chambers all yield a physical musical signal with much higher crest factors. Typical crest factors for musical instruments are on the order of 18 to 20 dB. That is, with musical instruments, peaks tend to be sharper than for speech. Both the TK-point of the compression detector, as well as the nature of the detector itself, needs to be different to prevent the hearing aid amplifier from entering compression prematurely. Clinical experience has shown that compression systems that use an RMS detector rather than a peak detector may be a more appropriate choice for music. If a peak detector were to be used, the compression kneepoint should be set 5- to 8-dB higher than for equivalent intensities of speech to prevent the music from forcing the hearing aid into its nonlinear mode prematurely.

2.5. Different Intensities for Speech and Music

Typical outputs for normal intensity speech can range from 53-dB sound pressure level (SPL) for the [th] as in *think* to about 77 dB SPL for the [a] in *father*. Shouted speech can reach 83 dB SPL. This 30-dB range (+12 to –18 dB) is related to the characteristics of the human vocal tract and vocal chords. Music can be on the order of 100 dB SPL with peaks and valleys in the spectrum of ± 18 dB. In fact, peaks for a 100 dB SPL musical input can cause conventional hearing aid micro-

phones to distort as the maximum microphone transduction capability is 115 dB SPL.

Clearly, for a given hearing loss, a musical input would require less gain to have the same output as a typically less intense speech input. Hearing aids need to be designed with unity gain above a certain input level to handle the lower gain requirements with musical input.

Looking at this from another point of view, many musical inputs are limited or clipped just after the microphone, prior to any amplification. Although this peak input-limiting level is quite adequate for speech, it is too low for typical musical inputs that have a much greater dynamic range than speech. As will be discussed in Section 4, the peak input-limiting level should be elevated to at least 105 dB SPL (and most probably 115 dB SPL) so that amplified music may retain its fidelity.

3. Music Parameters

Pitch is arguably the most important perceptual dimension in music. Variations in pitch are central to our experience of melody, harmony, and key. Individuals of the same culture with normal or corrected-to-normal hearing tend to experience melody, harmony and key in a similar manner—that is, sensitivity to these constructs does not require formal training (Bigand, 1993; Krumhansl, 1990).

Another important parameter in music is timbre. Timbre allows us to distinguish one instrument from another when both are playing the same note. The physical variables that contribute to our experience of timbre include the spectrum, temporal envelope, and transient components of a tone. Grey (1977) found that timbral distinctions are best described along three dimensions: spectral energy distribution (*ie,* bandwidth and concentration), synchronicity of the temporal envelope across partials, and onset characteristics (*eg,* speed of attack) (also see Krumhansl, 1989; McAdams *et al.,* 1995).

What follows is a series of proposals regarding the effects of distortion on pitch and timbre. These proposals are informed by the psychoacoustic and music cognition literature and may have implications for both the design and fitting of hearing-assistive devices. Proposals have been organized under the categories of linear and nonlinear distortions.

3.1. Linear Distortions

Linear distortions are defined here as changes in the intensity or phase of individual frequency components without the addition of new components. Because timbre is influenced by the spectral distribution of energy, an imbalance in the amplification of low- and high-frequency channels will always affect timbre. Having a consistent spectral distribution of energy over a series of tones serves to perceptually glue the tones together. If the tones of a melody possess spectra that modulate unpredictably, the coherence of a melody may deteriorate (Bregman, 1990). This sort of deterioration could prove to be particularly challenging for a musician who participates in a group performance where streaming demands are high. Thus, the balance of amplification in low- and high-frequency channels should remain consistent over time.

Modulating the amplification in low- and high-frequency channels may also lead to problems for musical pitch perception. For example, pitch discrimination thresholds are increased when a large change occurs in the spectral distribution of energy from one tone to the next (Warrier and Zatorre, 2002). Related to pitch discrimination is the musical construct known as interval size, that is the perceived distance between two pitches. Although listeners are quite good at detecting differences in interval sizes under normal circumstances, this ability breaks down substantially when the distribution of spectral energy modulates from tone to tone. For example, an ascending interval can be made to sound larger than it normally would if it involves a transition from a bass-weighted spectrum to a treble-weighted spectrum (Russo and Thompson, in press, 2004).

Historically, sound reproduction systems (including hearing aids) have been designed without much consideration for phase distortion. Indeed, the classic view in psychoacoustics had been that the auditory system was insensitive to phase (Helmholtz, 1877/1954). It is now known that listeners are sensitive to phase relationships, particularly for tones of low frequency and rich harmonic structure (Galembo *et al.,* 2001). Although phase distortion introduced by a stereo system in an open room will be relatively benign (sound reflections have the effect of randomizing phase), phase distortions introduced by hearing aids enter the ear canal directly and may be quite

audible as a result. If these distortions are binaural (varying across channels), the auditory image may be experienced as being in motion. If these distortions are monaural (consistent across channels), the experience of pitch may be altered.

3.2. Nonlinear Distortions

Nonlinear distortions involve the addition of new harmonics (normally of the fundamental) not present in the original signal. A common source of such distortions in hearing-assistive devices occurs when the amplitude of the signal is driven beyond the limits of the receiver, that is peak clipping. The addition of harmonics owing to peak clipping will almost always have negative consequences for timbre. Peak clipping may be particularly noticeable in instruments like the clarinet that produce tones with odd harmonics and instruments like the piano that produce inharmonic tones. In the case of a tone with odd harmonics, the addition of even harmonics would clearly alter the timbre. In the case of an inharmonic tone, new harmonics of the fundamental energy would lead to beating between harmonic and inharmonic partials. This beating would add roughness to the timbre as well as reduce pitch clarity. In contrast, the result of peak clipping may be less noticeable in a harmonic tone with odd and even harmonics, as the distortion products would introduce neither new partials nor beating but rather a shift in the spectral profile.

The standard alternative to peak clipping is compression. Although compression is necessary to deal with the wide dynamic range inherent in music, an overly active compression scheme could be problematic. Music is normally composed with a metrical grid in which prominent pitches occur regularly in time. These prominent pitches are primarily indicated by intensity. Too much compression will minimize intensity differences from tone to tone and, as a result, may impede perception of relations between prominent pitches. Moreover, intensity changes are commonly used to instantiate expectancy (Jones *et al.,* 2002) and convey emotion in speech as well as music (Sherer, 1986; Sloboda, 1991).

Clinical experience suggests that the compression systems for most forms of music should be set with a relatively low compression ratio (*eg,* 1.5:1) and with a relatively high TK setting (*eg,* 65–75 dB). If the hearing aid is vented, which may be the typical case for a mild-to-moderate hearing loss, the effective compression ratio would actually be less (perhaps 1.3:1) because unamplified sound would enter through the vent, bypassing the hearing aid electronic processing.

Another important form of distortion results when two or more partials interact within a nonlinear system (*ie,* intermodulation). The result of such interaction is the addition of partials that are inharmonic, or (*ie,* not harmonically related to the fundamental). Pitch models are unanimous in their prediction that the perception of pitch will be degraded for inharmonic tones (*eg,* Moore, 1997; Terhardt, 1974; Schouten *et al.,* 1962). In addition, sensitivity to tonal relations is degraded for pitch sequences composed of inharmonic tones (Russo *et al.,* 2000). The effect of inharmonicity is most audible for low-frequency tones below 100 Hz (Järveläinen *et al.,* 2001).

4. Hearing Aid Parameters

4.1. Peak Input-Limiting Level

Hard-of-hearing musicians have long complained about the poor sound quality while playing their instrument or when listening to music through hearing aids. Indeed, many nonmusicians also complain of reduced sound quality while listening to music through their personal amplification device. Traditional approaches by the hearing aid industry to optimize a hearing aid or a music program within a hearing aid are varied. Some manufacturers have sought to reduce the low-frequency amplification and output; others have sought to increase this gain and output. Other manufacturers have sought to increase mid-frequency gain and output to optimize the long-term spectrum of music. Still others have recommended that many of the noise reduction algorithms for a music input be disabled because they may incorrectly identify some music input (*eg,* flutes) as noise or feedback. These approaches have met with limited clinical success.

Because most engineers have traditionally had a speech input in mind when designing hearing aids, it is understandable that the peak input-limiting level of hearing aids has been set to about 85 dB SPL. That is, just after the microphone, a limiter exists that prevent inputs in excess of 85 dB SPL from being transduced through

the hearing aid. This is very reasonable for speech because the most intense components are less than 85 dB SPL. Anything of greater intensity than 85 dB SPL must therefore not be speech and should be limited. The peak input-limiting level is not reported on ANSI hearing aid specification sheets and usually conversations with someone from the engineering department of the hearing aid manufacturer are needed to determine what this level actually is for a particular hearing instrument. Yet, modern hearing aid microphones can readily transduce up to about 115 dB SPL with limited distortion.

Such an 85-dB peak input-limiter level design rationale works well for speech of all languages. The low-back vowel [a] is found in all languages, and regardless of language, no human speech sounds can be of greater intensity than [a]. However, this can cause a problem for nonspeech music inputs. The important question arises concerning the importance of elevating the peak input-limiting level above 85 dB SPL to a level that is more in line with those inputs typically found in music.

To study this issue, a wearable experimental hearing aid was constructed where the peak input-limiting level could be altered in discrete steps from 115 dB SPL, 105 dB SPL, 96 dB SPL, to 92 dB SPL. The gain, frequency response, and output of these four conditions were within 3 dB, regardless of peak input limiting level. This was originally reported in Chasin (2003).

Measures of distortion (signal-to-distortion ratio) and patient quality judgment scales were used. Distortion measurements can be affected by many electroacoustic parameters, especially with nonlinear devices. Cross-correlation (and autocorrelation) has been used widely in this subject area, but as pointed out by Kates and colleagues (1992, 1994), this only is valid if the amplification scheme is linear. Kates (2000) demonstrated that for nonlinear hearing aids, a notch paradigm works better, specifically showing that "debris or filling-in" of a spectrum with a well-defined notch could be used as a measure of fidelity. If there is no difference (or "fill-in") within the notch between the input and the output of a hearing aid, then this could be construed as perfect fidelity. While Kates (1992) suggested the notches found in a comb filter, such notches can also be created using clinically accessible software such as Adobe Audition (formerly called Cool Edit) (Adobe Systems, Inc, Seattle, WA). The signal-to-distortion ratio can then be measured. In this case, a 0-dB signal-to-distortion ratio means no fill-in of the notch in the output spectrum, implying perfect fidelity. This filtered notch, centered in the 2000- to 3000-Hz region is shown in Figure 1. The 2000- to 3000-Hz region that was selected was purely arbitrary, however there is no reason to suspect differing results for other mid-frequency bands.

Measures of sound quality were obtained by using five, five-point perceptual scales that are relevant to music. This is a modification of the work of Gabrielsson and colleagues (1974, 1991) that has been used extensively in the hearing aid industry (see for example, Cox and Alexander [1983]). Specifically, hard-of-hearing musicians were asked to rate from 1 (poorest) to 5 (best) the following perceptual scales: Loudness, Fullness, Crispness, Naturalness, and Overall Fidelity. A perfect perceptual reproduction score was 25 points (5 × 5 scales).

Subjects were given the following definitions of the five perceptual parameters:

- Loudness: the music is sufficiently loud, in contrast to soft or faint;
- Fullness: the music is full, in contrast to thin;

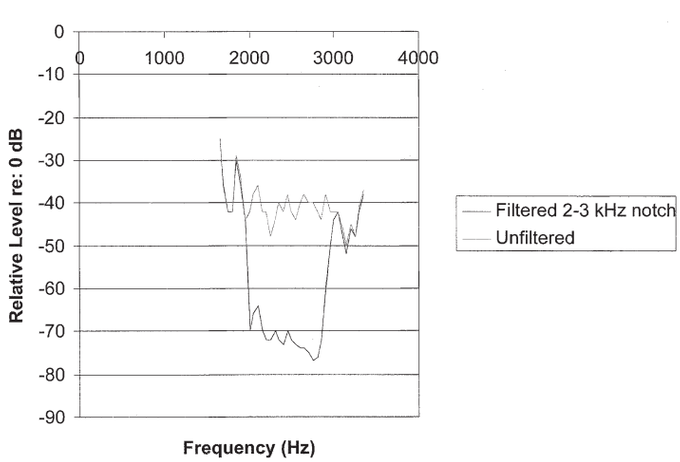

Figure 1. Unfiltered and filtered (between 2000 and 3000 Hz) notch from an intense music spectrum. If the output-filtered notch is similar to this input-filtered notch, then this is evidence of good fidelity or distortion-free hearing aid processing. This figure is reprinted from Chasin M (2003) with the permission of the publisher.

- Crispness: the music is clear and distinct, in contrast to blurred, and diffuse;
- Naturalness: the music seems to be as if there is no hearing aid, and the music sounds as "I remember it;" and
- Overall Fidelity: the dynamics and range of the music is not constrained or narrow.

Measures of signal-to-distortion ratio and measures of sound quality were assessed in 53 professional musicians (37 men, 16 women) who had music induced and/or presbycusic hearing losses. The age range of the musicians was from 33 to 81 years.

Figures 2 and 3 show the fill-in in the square notch as the peak input-limiting level was dropped from 115 dB SPL, to 105 dB SPL, to 96 dB SPL, and finally to 92 dB SPL. For clarity, Figure 2 shows the data for 115 dB SPL and 105 dB SPL, whereas the data for the 96 dB SPL and 92 dB SPL conditions are shown in Figure 3.

Quantifying Figures 2 and 3, Table 1 shows the signal-to-distortion values that were measured in a 2-cc coupler, but data were also measured in the real ear for all 53 subjects. To hear the difference of decreasing the peak input-limiting level from 115 dB SPL successively down to 92 dB SPL, wave files have been generated and can be found at *http://www.randomizedtimer.net/musicandhearingaids* for both average conversational speech (65 dB SPL) and typically intense music (90 dBA and 100 dBA).

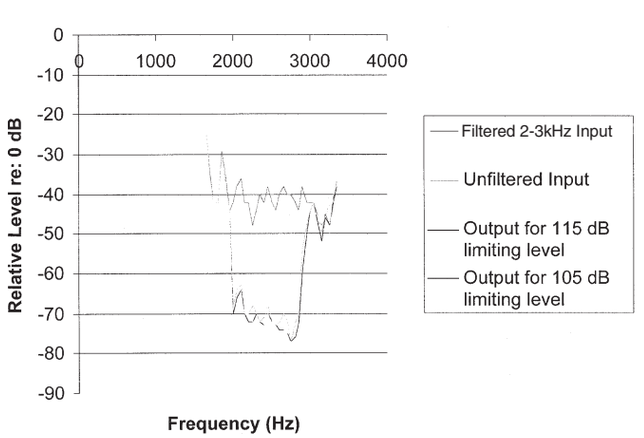

Figure 2. With peak input limiting levels of 115 dB SPL and 105 dB SPL, minimal "debris" can be seen in the filtered notch, suggesting good fidelity with minimal distortion. This figure is reprinted from Chasin (2003) with the permission of the publisher.

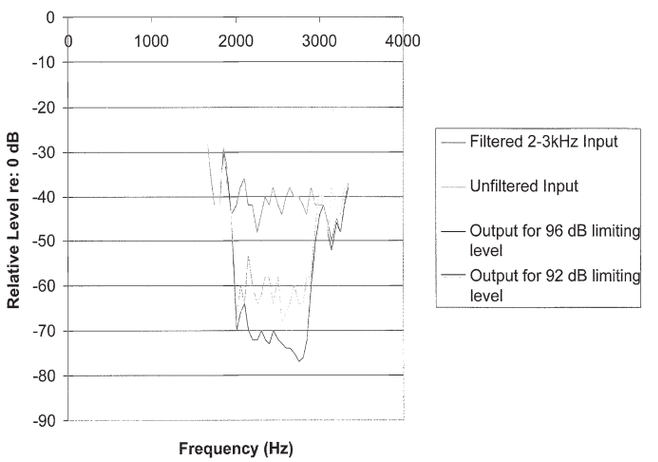

Figure 3. With peak input-limiting levels of 96 dB SPL and 92 dB SPL, "debris" can be seen in the filtered notch, suggesting poor fidelity with significant hearing aid distortion. This figure is reprinted from Chasin (2003) with the permission of the publisher.

Figure 4 shows the sum of the five perceptual scales (maximum value of 25) plotted against the measured signal-to-distortion ratios for the four peak input-limiting levels.

Statistically, the sum total of the five, five-point perceptual attribute scales for peak input-limiting levels of 115 dB SPL and 105 dB SPL

Table 1. Values for the Signal-to-Distortion Ratios from Figures 2 and 3*

Peak Input Limiting Level (dB SPL)	Signal-to-Distortion Ratio (dB)
115	–0.75
105	–2.25
96	–10.00
92	–10.5

*Perfect fidelity with no hearing aid processing distortion would be 0 dB. The more negative the number, the poorer the reproduction of music through the hearing aid.

(p = .08) did not differ, and no statistical difference was noted for the two lower-peak input-limiting levels of 96 dB SPL and 92 dB SPL (p = .12). However, a statistical difference was evident between the upper two levels and the lower two levels ($p < .001$).

In addition to the above measures, all subjects preferred the 115 dB SPL and 105 dB SPL levels anecdotally. Some comments were: "Music was so much more natural at the higher peak input limiting levels;" "My horn sounded like it should at the higher levels;" and "While conducting, I could hear the orchestra beautifully." Still, one of the 53 subjects noted that even at the highest input limiting level, "occasionally the drums and a couple of notes on the trumpet sounded odd."

Part of the appeal of loud music may be its ability to stimulate the vestibular as well as auditory system. Todd and Cody (2000) found that the sacculus—a primitive mechanism found in fish and humans—responds to music that is played above 90 dB SPL, particularly for music composed predominantly of low-frequency energy. Although it may be hard to fathom what benefit sound stimulation of the vestibular system would have for humans, the benefit to our fishy ancestors is less mysterious. Also of interest, the sacculus is known to have connections to the hypothalamus, a part of the limbic system responsible for drives such as hunger and sex.

Clinically, one valid suggestion for a client who is a musician or who likes to listen to music and who is fitted with a less-than-optimal hearing aid would be to reduce the input to the hearing aid (*eg,* turn down the stereo or Walkman) and turn up the hearing aid volume to reestablish the desired output. This reduction in input is like ducking under the low overhang and then standing upright after.

Clinical intervention can also take the form of using assistive listening devices. Depending on the manufacturer, the direct audio input or inductive coupling from an FM or an infrared system may be able to bypass the peak input limiter and couple directly into the unrestrained amplification stages in the hearing aid. However, an informal poll indicates that the majority of hearing aid manufacturers wire the direct audio input prior to the peak input limiter, whereas a minority do not. Users of devices from hearing aid manufacturers that wire into the direct audio input after the peak input limiter would benefit from the use of an assistive listening device for listening to music, as long as the hearing aid microphone is disabled.

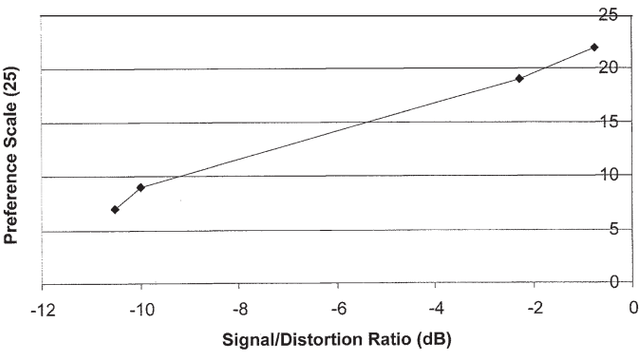

Figure 4. Preference scale results for the four levels of peak input limiting level (and associated signal-to-distortion ratios) for 53 musicians. Best sound quality was judged for peak input limiting levels of 115 dB SPL and for 105 dB SPL, while poor quality values were obtained for the lower levels. This figure is adapted from Chasin (2003) with the permission of the publisher.

A third clinical option if a person is already fitted with a less-than-optimal hearing aid because of a low peak input-limiting level is to use a well-defined attenuator placed over the hearing aid microphone. This would reduce the input effectively by the amount of the attenuation, thereby providing additional "head room." The music program on the hearing aid should be set to have an increased gain and output to reestablish the output in the ear canal. The music program should then only be used when the attenuator is placed over the hearing aid microphone. An example of such an attenuator is shown in Figure 5, where an Adhear wax guard (Hearing Components Inc., North Oakdale, MN) brushed with Whiteout (Paper Mate, Sanford Corp, Bellwood, IL) was placed over the two directional microphone ports of an in-the-ear digital hearing aid. Normally the Adhear strips are acoustically transparent, but when brushed with Whiteout, they provide 10 to 15 dB of uniform attenuation, from 750 Hz to about 6000 Hz. Ten to 15 dB of gain and output was added to the program to compensate for this. This brute-force method worked for the individual in question. When he obtains his next hearing aids, this would not be necessary.

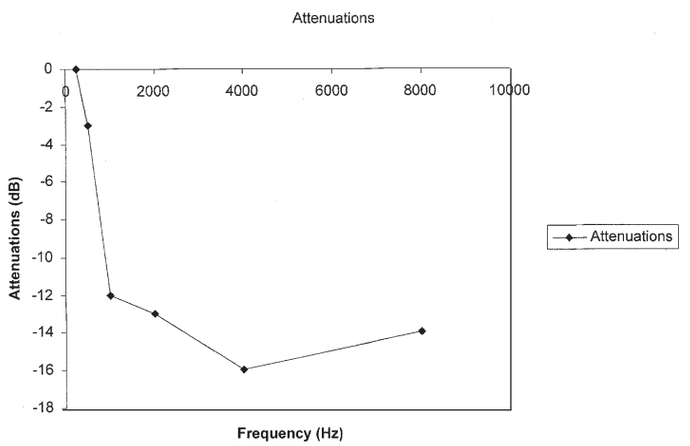

Figure 5. A brute force method: Attenuations caused by placing Adhear wax guards, painted lightly with Whiteout, over the hearing aid microphone ports to lessen the effective input to the hearing aid for a "music program". The music program should have this much gain and output added in order to maintain equivalent output.

4.2. Three Other Electro-Acoustic Parameters

Clearly the peak input limiting level is a major determining factor in the optimal reproduction of music, but three other more secondary factors are also necessary.

4.2.1. *Parameter 1*

For music, one channel (or a multichannel device where the gain of each channel is set at a similar level) is probably the best. Unlike speech, music requires that the relationship or balance between the lower-frequency fundamental energy and the higher-frequency harmonic energy remains intact to achieve optimal sound quality. This is especially true for violinists and violists as well as hard-of-hearing listeners who like to listen to a lot of classical music. Use of a single-channel (or similar gain in each channel) hearing aid that maintains this balance is therefore necessary. For woodwinds and quieter music, or if there is a precipitous audiometric configuration, a multichannel hearing instrument with different gain settings may be acceptable. If multichannel hearing aids are to be used, caution should be exercised to ensure that the adjacent channel compressors are set to be not too dissimilar.

4.2.2. *Parameter 2*

The knee-point on the input compression circuit should be set at approximately 5- to 8-dB higher for music than for speech. Speech has a crest factor of about 12 dB. The crest factor is the difference between the RMS of the signal and the peak. Because the human vocal tract is so inherently damped (soft lips, nasal cavity, soft palate, etc.), the peaks of speech are closer to the average RMS. In contrast, because music is generated from electrical or hard-walled cavity instruments, less inherent damping occurs. The difference between the average RMS and the peak levels in a music spectrum is therefore greater. Typical crest factors for music are on the order of 18 to 20 dB. Subsequently, the compression circuit should not be allowed to enter its nonlinear phase prematurely. Having a higher compression knee-point is one of several methods of accomplishing this. This would be more of an issue for peak compression detectors and would be less important for those hearing aids that use an RMS-based compression detector.

4.2.3. *Parameter 3*

A good argument can be made for a modified wide dynamic range compression (WDRC) circuit for musicians. Most musicians, especially if they suffer from music exposure and/or presbycusis, have a mild-to-moderate hearing loss with poorer acuity in the mid- to high-frequency ranges. This type of hearing loss is predominantly outer hair-cell damage, and so the WDRC circuit is typically optimal. However, the slightly higher settings of the knee-point means, strictly speaking, that the circuit would not be wide dynamic range compression, but a more narrow or high-level version of this circuit.

An interesting study would be for musicians to be able to listen to (or play) various types of music and give them control over various compression parameters. We would expect the data-logged results to support the WDRC view.

Given the characteristics of today's hearing aid technology and the types of input spectra that musicians and nonmusicians who like to listen to music are subject to, the peak input-limiting level should be at least 105 dB SPL. The exact setting of this level probably has to do with the type of music as well as the preferred instruments that the musician plays. Almost all hearing aids on the market have a restricted peak input-limiting level (typically 85–90 dB SPL). Modern microphones

can safely transduce 115 dB SPL without appreciable distortion, so there are few engineering or audiologic reasons to limit the input range of today's hearing aids. Of course, the output can and should be limited, depending on the gain requirements of the individual's hearing loss.

The problems of a peak input-limiting level that is too low are independent of the processing method, which means this is as much a problem for analog hearing aids as it is for digital. And because the peak input limiting level is not disclosed on manufacturer specification sheets, it is up to the hearing health care professional to obtain this information, either from the manufacturer's representative or from one of its engineers, prior to selecting amplification for a hard-of-hearing musician or someone who likes to listen to music.

The optimal hearing aid for musicians and those with hearing losses who like to listen to music includes a high peak input-limiting level of at least 105 dB SPL, WDRC with a higher TK-point than prescribed for speech, and for most musicians, a one-channel (or multichannel system with similar compression specifications) system. A popular but dated example of one such hearing aid that comprises elements of all four parameters is the K-AMP circuit (Etymotic Research, Inc, Elk Grove Village, IL).

Other technologies can simulate a wider input range of sounds to the hearing aids other than by elevating the peak input-limiting level, and these should be assessed clinically. Some of these technologies involve compressing the input prior to the peak input limiter, followed by a form of expansion after this limiter—it is like ducking under a low hanging overhead or flying under a bridge. Merely altering the frequency response of the hearing aid will not be beneficial for musicians and those music listeners who are not musicians. The input spectrum of music is inherently quite variable and certainly different than that of speech. Hearing aids need to be designed with this in mind from the very onset of development.

4.3. Feedback and Noise Reduction Systems

Many modern digital hearing aids have features that seek to minimize environmental noise and the propensity for acoustical feedback. These have been shown to be useful in many situations but may have some drawbacks for some types of music. For example, clinically, some feedback reduction systems erroneously target the flute, resulting in an unwanted reduction in this instrument's intensity. Depending on the algorithm, other systems have no difficulty with this instrument. In contrast, some feedback reduction systems are designed to minimize "feedback-like" sound just after the hearing aid microphone, thereby reducing the input such that an overly low peak input-limiting level ceases to become a problem.

However, to give a general comment would be misleading at this point. Different manufacturers use different algorithms to accomplish feedback reduction and noise reduction. One cannot definitively say that a particular feature should be disabled for a music program. Many manufacturers do state that the music program of their digital hearing aid should be as plain as possible, without any of the features available for speech input programs. This may or may not be the case for music, and clearly, more research is required in this area.

5. In-Ear Monitors

For the past decade, personal in-ear monitors have been available for the musician. These have grown out of a need to control and minimize the music exposure up on stage and were designed to replace the wedge speakers and other stage-based monitor systems. The rationale is that if an optimal sound mix can be achieved in the ears of the musician, the overall result will be less damaging exposure and an improved quality of sound.

Traditionally, guitarists needed their amplification to be higher because the lower-frequency bass or keyboard tones would drown them out. A tug-of-war ensued, with each musician fighting against the other's sound levels. This resulted in a much higher level than what was required, with poor balance between the various instruments and vocalists in the band.

In-ear monitors neutralize the environment. Individual musicians can have their own mix and their selected level. If properly fit and mixed, guitar players may have adequate monitoring of their own instruments, that of the others in the band, and even some input from the audience microphone. In short, ear monitors ensure a proper balance of music and monitoring, usually at a much lower level.

In-ear monitors can be either custom made (see Figure 6) or one-size-fits all. They are similar to hearing aids in that there is a preamplifier and usually one or two receivers (with associated crossover network). The microphone is replaced by a direct audio input from the sound engineer's rack. A cable connects the ear monitors either directly to the rack, such as in the case of a drummer who does not need to move around, or via a wireless FM transceiver that allows mobility. Optimal settings of the musical mix are typically determined by the sound engineer in conjunction with the musician. Optimal equalizer levels are set with the assistance of the hearing health care professional. It is the hearing health care professional's responsibility to provide the musician with a frequency-by-frequency listing of the corrections necessary to obtain a flat response.

In addition, if a hearing loss has occurred, level-dependent corrections should also be provided such that the musician can wear the device as a hearing aid (with music input). The level-dependent and frequency-response specifications for hard-of-hearing musicians are obtained in the same way as that of any hearing aid prescription. Ensuring a flat frequency response in the ear of the musician requires some extra calculation.

The calibration technique to ensure a flat response involves the use of a real-ear measurement system and a flat noise source such as a good quality white noise. Some audiometers generate a fairly flat white noise whereas others do not. A number of software programs—the Adobe Audition is one such software program, and most are relatively inexpensive—can generate a perfectly flat noise source and this could be recorded onto tape or CD and fed directly in to the ear monitor. The input cable for the in-ear monitor can be plugged in to this white noise source via the audiometer or computer, and a white noise input can be generated at about a 70-dB audiometer dial reading. A probe tube can be placed in between the in-ear monitor and the ear canal wall in the normal fashion, and the output in that individual's ear can be measured. Deviations from a flat response can be noted and provided to the musician to share with the sound engineer. For example, if there was an 8-dB resonance at 2000 Hz, the sound engineer could attenuate this frequency region by 8 dB.

The real-ear measurement system should first be configured to receive external stimuli. This can be accomplished by disabling both the reference microphone and the speaker. Various manufacturers of real-ear measurement systems have different methods of accomplishing this, but they usually involve setting the stimulus level to 0 dB (*eg,* Audioscan, Dorchester, Ontario Canada) or turning the reference microphone to *off* (*eg,* Frye Electronics, Tigard, OR). Performing this 2-minute calibration ensures that any subsequent changes to the music are because of the musician's personal preferences rather than any limitations of the in-ear monitors.

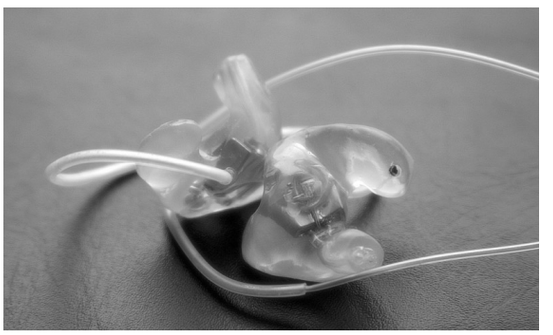

Figure 6. Example of a custom made ear monitor with associated cables. Photograph courtesy of Sensaphonics. Used with permission.

References

Bigand E. The influence of implicit harmony, rhythm, and musical training on the abstraction of "tension-relaxation schemas" in tonal musical phrases. Contemporary Music Review 9:123-138, 1993.

Bregman AS. Auditory scene analysis. Cambridge, MA: MIT Press, 1990.

Chasin M. Music and hearing aids. Hearing Journal 56:7, July 2003.

Cox RM, Alexander GC. Acoustic versus electronic modifications of hearing aid low-frequency output. Ear Hear 4(4):190-196, 1983.

Gabrielsson A, Rosenberg U, Sjogren H. Judgments and dimension analyses of perceived sound quality of sound-reproducing systems. J Acoust Soc Am 55:854-861, 1974.

Gabrielsson A, Hagerman B, Bech-Kristensen T. Perceived sound quality of reproductions with different sound levels. Report TA No. 123, Karolinska Institutet: Stockhom, ISSN 0280-6819, 1991.

Galembo A, Askenfelt A, Cuddy LL, Russo FA. Effects of relative phases on pitch and timbre in the piano bass range. J Acoust Soc Am 110:1649-1666, 2001.

Grey JM. Multidimensional perceptual scaling of musical timbres. J Acoust Soc Am 61:1270-1277, 1977.

Helmholtz H. On the sensations of tone. New York: Dover, 1877/1954.

Järveläinen H, Välimäki V, Karjalainen M. Audibility of the timbral effects of inharmonicity in stringed instrument tones. Acoustics Research Letters Online 2:79-84, 2001.

Jones MR, Moynihan H, MacKenzie N, Puente J. Temporal aspects of stimulus-driven attending in dynamic arrays. Psychol Sci 13:313-319, 2002.

Kates JM. On using coherence to measure distortion in hearing aid aids. J Acoust Soc Am 91:2236-2244, 1992.

Kates JM. Cross-correlation procedures for measuring noise and distortion in AGC hearing aids. J Acoust Soc Am 107(6):3407-3414, 2000.

Kates JM, Kozma-Spytek L. Quality ratings for frequency-shaped peak-clipped speech. J Acoust Soc Am 95(6):3586-3594, 1994.

Kent RD, Read C. Acoustic Analysis of Speech (2nd ed.), New York: Delmar, 2002.

Krumhansl CL. Why is musical timbre so hard to understand? In J. Nielzen O. Olsson (eds.): Structure and Perception of Electroacoustic Sound and Music. Amsterdam: Elsevier, Excerpta Medica, pp. 45-53, 1989.

Krumhsansl CL. Cognitive foundations of musical pitch. Oxford: Oxford University Press, 1990.

McAdams S, Winsberg S, Donnadieu S, et al. Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. Psychol Res 58:177-192, 1995.

Moore BCJ. An Introduction to the Psychology of Hearing (4th ed). London: Academic Press, 1997.

Russo FA, Galembo A, Cuddy LL. Variability in musical pitch perception across the tessitura of the piano. Paper presented at the Society for Music Perception and Cognition, Toronto, Canada. Published in Musical Intersections, 2000, 256, 2000.

Russo FA, Thompson WF. An interval-size illusion: The influence of timbre on the perceived size of melodic intervals. Perception Psychophysics, 2004 (in press).

Scherer KR. Vocal affect expression: A review and a model for future research. Psychol Bull 99:143-165, 1986.

Schouten JF, Ritsma RJ, Cardozo BL. Pitch of the residue. J Acoust Soc Am 34:1418-1424, 1962.

Sloboda J. Music structure and emotional response: Some empirical findings. Psychol Music 19:110-120, 1991.

Terhardt E. Pitch, consonance and harmony. J Acoust Soc Am 55:1061-1069, 1974.

Todd NP, Cody FW. Vestibular responses to loud dance music: A physiological basis of the “rock and roll threshold”? J Acoust Soc Am 107(1):496-500, 2000.

Warrier CM, Zatorre RJ. Influence of tonal context and timbral variation on pitch perception. Percept Psychophys 64:198-207, 2002.