

## Application of Psychoacoustics to Audio Signal Processing

Brent Edwards, Sound ID, Palo Alto, CA

### Abstract

*The application of psychoacoustics to audio signal processing has been successfully applied to bit-rate reduction coding schemes such as MPEG but have not seen significant implementations into other audio applications. Because of the nonlinear aspect of auditory perception, which includes peripheral mechanics and central cognitive abilities, algorithms derived to enhance one aspect of hearing may have a completely unexpected consequence on other hearing abilities. This paper discusses the use of basic psychoacoustics for both signal processing design and validation. Since the field of hearing aid technology has broadly applied basic psychoacoustics to hearing aid design, it will be used to describe how these measures can be used in signal processing design and validation.*

### 1: Introduction

For the majority of the past century, sound production models have been far more prevalent than sound perception models in signal processing applications. Speech production models are the basis of vocoders used in the telecommunications industry and have been essential in speech recognition applications. Production models are obviously also essential in music and speech synthesis algorithms.

The application of perception-based models, and psychoacoustics in general, to signal processing design has increased significantly over the past twelve years, driven by the prevalence of digital multimedia systems in consumer entertainment products. Perceptual-coding has been incorporated in bit-rate reduction algorithms such as MP3 and ATRAC. These rely on masking patterns that primarily result from the excitation representation of sound in the cochlea. Virtual 3D audio algorithms manipulate interaural time and intensity differences and spectral shapes (through Head-Related Transfer Functions) in order to produce a more realistic perception of an auditory field. Speech recognition systems have also used auditory models as front-ends in order to emulate the pre-processing produced before the brain does its own speech recognition.

Most of these perceptually-based signal processing schemes were designed to take into account a single aspect of auditory perception, e.g., masking or localization. The auditory system is highly nonlinear, however, and processing that alters one aspect of perception can have a deleterious effect on other aspects. Processing to enhance speech formants can affect masking and suppression effects. Manipulations of binaural signals can alter the auditory system's ability to suppress the perception of echoes.

Usually, signal processing designs that do take psychoacoustics into account only include one aspect of auditory perception, ignoring all of the other auditory abilities that may be affected by the processing. The application of more than the primary psychoacoustic ability of interest to the signal processing design can be difficult. Less difficult, however, is to investigate secondary perceptual effects during the signal processing validation stage, after the design has been implemented. Here, basic psychoacoustic tasks can be performed on processed stimuli in order to determine the effect of the signal processing on the areas of auditory perception that were not directly considered in the signal processing design. For example, echo perception could be investigated with a system that was designed to do beamforming on a conference telephone system. In this way, abnormalities in the perception of the signal processing can be parsed into specific categories of auditory areas and possibly addressed in future signal processing research. Relying simply on sound quality measures such as the Mean Opinion Score may detect that some part of the signal processing is not optimal but will not determine why. Electroacoustic measurements are also incomplete since they characterize the acoustic signal but not the perceptual one; distortion that is easily measured may not be actually be heard, while other electroacoustically characterized behavior may have unknown perceptual consequences.

The psychoacoustic characterization of hearing impairment and of normal hearing have always provided the primary design criterion for hearing aids. The standard audiogram which diagnoses hearing loss is the most basic of psychoacoustics tasks: the detection threshold level of a pure tone. As signal processing has become more sophisticated, so too have the application of

psychoacoustics to the design of hearing aids, to the point where cognitive aspects of hearing such as long-term adaptation have been applied to hearing aid design, where the signal processing adapts over time in order for the user to acclimatize to the new sounds. Because the hearing aid field has the most depth of psychoacoustic application to signal processing, the rest of this paper will focus on its use of psychoacoustics for both signal processing design and validation.

## 2: Application to hearing aids

The field of hearing aids stands out among the various audio processing fields because the signal processing can have very specific design goals derived from data on auditory perception. Ultimately, the goal of hearing aid design is to give the user of the aid the same hearing ability and function as someone with normal hearing—the perception of sound through a hearing aid by the hearing impaired should be the same as the perception of sound by normal hearing individuals.

### 2.1: Design specification

While hearing aids for most of the twentieth century were simple linear processors, psychoacoustic characterization of hearing loss had indicated for decades that this design was inadequate. Steinberg and Gardner [1] measured the abnormal growth of loudness in subjects with sensorineural hearing loss, noting that sensitivity to sound is more impaired at low levels than at high levels.

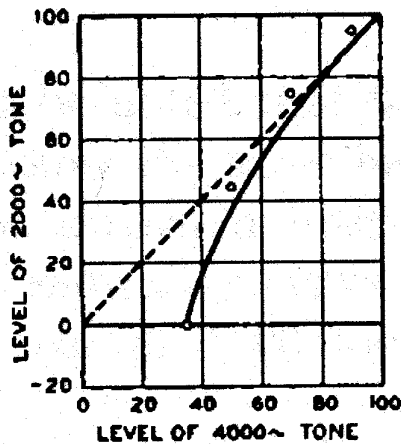


Figure 1. Matched loudness levels for a tone in a region of hearing loss (4 kHz) versus a tone in a region of normal hearing (2 kHz). The dashed line is when hearing at 4 kHz is normal. From [1].

Figure 1 shows this effect, plotting equal loudness of two tones where one is in a region of hearing loss and the other is not. The dashed line shows the expected data if both tones were in normal hearing regions. As can be seen, hearing aid gain must be reduced as sound level increases in order to restore loudness to normal. This produces a compressive effect in the input-output function since the gain reduction with level compresses the output dynamic range.

Figure 2 shows a typical input-output function for a hearing aid designed to restore loudness to normal and is similar to the functioning of the first commercial compression hearing aid that was introduced in 1989. The dashed line represents unity gain. This design was validated in 1991 with velocity measurements of the basilar membrane in chinchillas [2], which demonstrated the same difference between normal and impaired response to a pure tone physiologically as was found psychoacoustically over 50 years prior. This data is replotted in Figure 3. Hearing aid design specifications taken from this physiological data would have produced very similar input-output curves to that shown in Figure 2, indicating that psychoacoustic characterization can lead to signal processing that mimics cochlear mechanics.

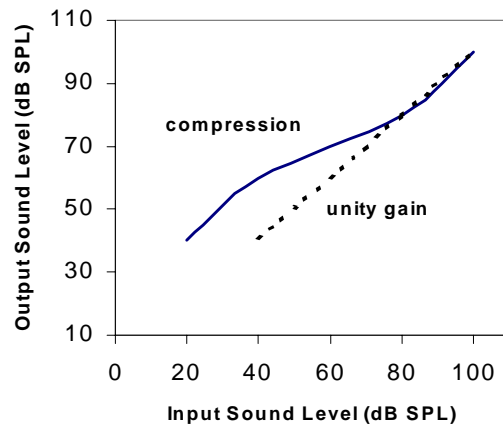
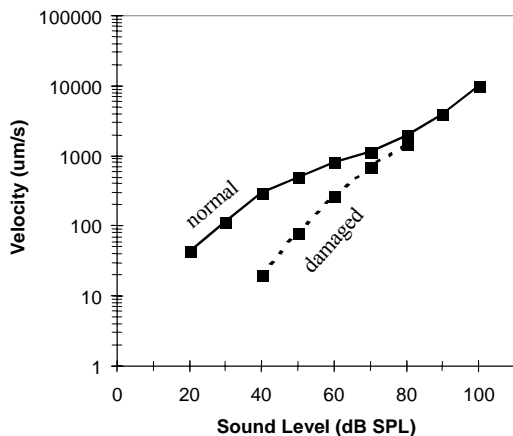


Figure 2. Input-Output function of a compression hearing aid (solid line). Unity gain is shown for reference (dashed line).

Envelope perception is an important aspect of hearing aid design since the envelope of speech carries significant intelligibility information [3]. While the I/O curve in Figure 2 tells us something about how envelopes will be processed and the perceptual I/O functions of Figures 1 and 3 tells us something about how envelopes will be perceived, they do not provide sufficient information for useful characterization of aided envelope perception because they lack dynamic information. Both compression

hearing aids and the auditory system have time constants associated with how quickly the gain is turned up and down in response to changes in stimulus level. In order to ensure that a hearing aid is designed to process envelopes properly, the dynamics of the auditory system must be known.

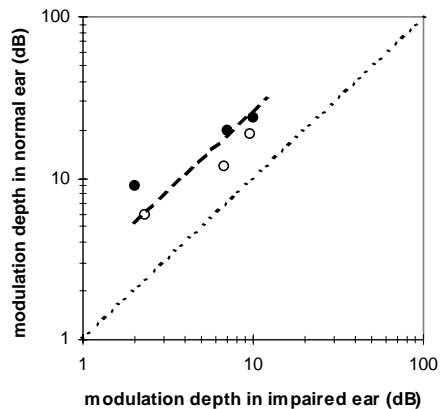


**Figure 3. Input-Output function of a healthy cochlea (solid) and one with hearing impairment (dashed) [2].**

Figure 4 shows data from a 16 Hz amplitude modulation matching psychoacoustic experiment using subjects who had normal hearing in one ear and impaired hearing in the other [4]. Subjects were asked to adjust the modulation depth of the sound in one ear until the envelope fluctuations were perceived to have equal strength in both ears. In other words, at the point of final adjustment, the modulation in the normal ear was perceived to have the same strength as in the impaired, even though the modulation depth in the two acoustic signals could be very different.

The data points show the adjusted levels of equally perceived envelope fluctuation strengths. The data indicates that envelopes need to be about 1/3 the strength in impaired ears in order to sound the same as in normal ears. This is consistent with the data in Figures 1 and 3 and results in the same compression design as in Figure 2. What this data also specifies is that the time constants of the hearing aid must be fast enough to effectively reduce the dynamic range of the 16 Hz envelope fluctuations. The time constants must be significantly faster than half the period of the envelope, or much less than 31.25 ms. If the time constants were longer, then the envelope would not be affected by the compression and the modulation depth at the output of the hearing aid would not be significantly different than at the input. Ideally, the envelope at the output of the hearing aid would have level fluctuations that

were 1/3 of the fluctuations at the input in order to restore the psychoacoustic distortion shown in figure 4 to normal.



**Figure 4. Points of perceived equal envelope fluctuation strength [4].**

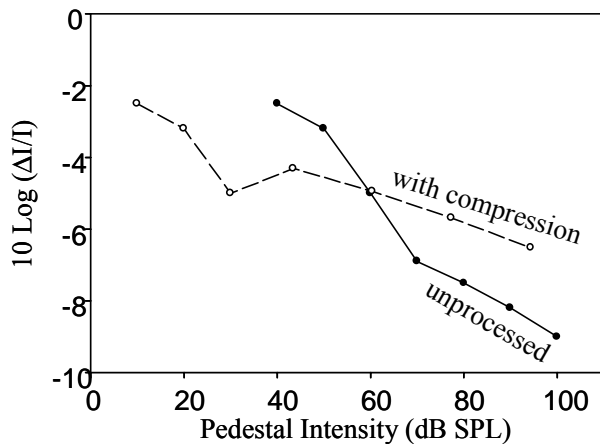
## 2.2: Design validation

Due to the complexity of the auditory system, signal processing to enhance one aspect of listening may have a negative effect on another aspect of listening. A system for enhancing speech formants may effect the envelope of speech, providing improvement in one aspect of speech while deteriorating another. The fundamental basis of most hearing aid signal processing is intended to restore certain psychoacoustic aspects of hearing to normal, such as loudness, but this may affect other aspects of hearing in unknown ways. Additional signal processing is typically added to digital aids such as noise reduction, directional processing and feedback reduction algorithms. All of these affect basic perceptual abilities. In order to determine what effects these have on perception, several basic psychoacoustic tasks should be analyzed and ideally measured with subjects listening to sound processing by the signal processing algorithms. Currently, algorithms are validated using global perceptual metrics, such as speech recognition scores and sound quality ratings. By using basic psychoacoustic metrics, a more fundamental understanding of the result of the signal processing can be understood.

Intensity discrimination is a task that measures sensitivity to level by determining the threshold at which a listener can detect a difference in level of a sound. Hearing aids are not designed with any attempt to normalize this ability. In fact, several criticisms of compression suggest that it makes intensity discrimination more difficult by reducing intensity differences between sounds; a 3:1 compression ratio could reduce a 6 dB level difference to 2

dB. Therefore, it is important to know whether compression in a hearing aid does affect the ability to detect level differences.

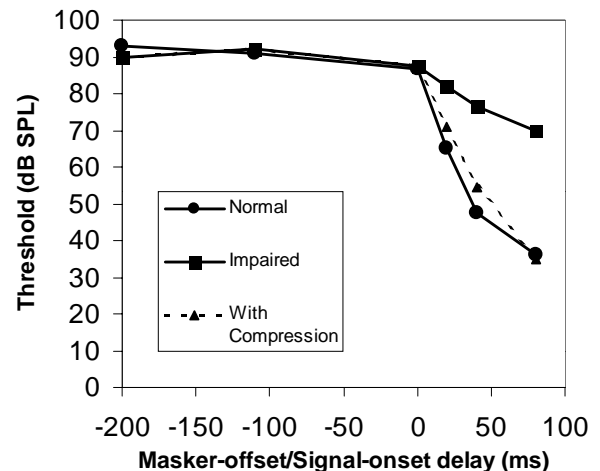
The solid line in figure 5 plots discrimination performance with normal hearing listeners [5]. Their ability to detect differences improves with level, which is why the line has a negative slope. Discrimination performance is plotted as a Weber fraction, or the ratio between the increment ( $\Delta I$ ) and the level of the sound on which the increment is added ( $I$ ). Since hearing aids add gain in addition to compression, the operating point of the discrimination task is moved to the right during the discrimination task, where the ability to discriminate is better. The dashed line shows the effective Weber fraction for intensity discrimination for a hearing impaired person listening to a compression hearing aid [6]. The results show that the hearing aid listener actually performs better than normal with the hearing aid for signals below 60 dB SPL and worse for levels above. This psychoacoustic characterization of perceptual behavior with the compression hearing aid tells how intensity coding has been affected by the combination of the hearing aid with their impairment, and could be used to explain whatever speech understanding and quality results are obtained with the aid. They also give guidance to improving the compression aid in future designs.



**Figure 5. Intensity discrimination, where the solid line shows normal performance for both hearing impaired and normal hearing [5]. The dashed line shows performance of the hearing impaired with a compression hearing aid [6].**

Forward masking is a measure of temporal resolution, measuring the ability of a listener to hear one sound in the presence of another when the two have similar spectra but are separated in time. The further the separation between the masking sound and the signal that listener has to detect, the lower the detection threshold of the listener.

People with hearing impairment show elevated forward masking thresholds, resulting in a poorer ability to separate sound. Two of the lines in Figure 6 show data obtained with hearing impaired and normal listening listeners [7]. A compression hearing aid can affect this measure since the aid will provide level-dependent gain, providing more gain to the signal than the masker if the time constants are fast enough to adjust between the masker and signal. If the time constants of the compressor are slow (greater than 100 ms), there will be no change to the shape of the impaired forward masking performance since the gain will not change from the masker offset to the signal onset. The dashed line shows the forward masking performance that is predicted from a fast-acting compression hearing, one in which the gain can change on the order of milliseconds. While such a compressor could be designed solely to restore abnormal loudness perception shown in Figure 1, the result shows that this compressor can restore forward masking ability to near normal, validating the compression design from the psychoacoustic domain of temporal processing ability.



**Figure 6. Forward masking data, plotting detection threshold of a brief signal as a function of temporal separation between the masker offset and signal onset. Plotted with solid lines are data from normal hearing and hearing impaired from [7], and a prediction of performance by the hearing impaired with a fast-acting compression hearing aid.**

### 3: Conclusions

Psychoacoustics can provide more detailed design criterion for audio signal processing than is currently being used. As demonstrated with the hearing aid discussion, this

can be done to provide accurate signal processing functioning to achieve the audio effect desired.

In addition, psychoacoustics can be used to validate signal processing design by measuring basic psychoacoustic characteristics with post-processed signals. Such measures can determine how the audio processing will affect basic psychoacoustic abilities and can detect abnormal perceptual behavior resulting from the signal processing but in perceptual categories that are not directly related to the primary target of the signal processing. Examples of secondary psychoacoustic effects that might not be part of the signal processing design criterion could be loudness, echo perception, temporal and spectral modulation detection and discrimination, suppression and many cognitive effects. Measuring the effect of the signal processing in these psychoacoustic areas can verify the proper performance of the signal processing with respect to auditory perception and lead to improved signal processing designs.

#### **4: References**

- [1] J. C. Steinberg and M. B. Gardner, "The dependency of hearing impairment on sound intensity," *J. Acoust. Soc. Am.*, vol. 9, pp. 11-23, 1937.
- [2] M. A. Ruggero and N. C. Rich, "Furosemide alters organ of Corti mechanics: evidence for feedback of outer hair cells upon basilar membrane," *J. Neurosci.*, vol. 11, pp. 1057-1067, 1991.
- [3] D. J. Van Tasell, S. D. Soli, V. M. Kirby and G. P. Widen, "Speech waveform envelope cues for consonant recognition," *J. Acoust. Soc. Am.*, vol. 82, pp. 1152-1161, 1987.
- [4] B. C. J. Moore, M. Wojtczak and D. A. Vickers, "Effects of loudness recruitment on the perception of amplitude modulation," *J. Acoust. Soc. Am.*, vol. 100, pp. 481-489, 1996.
- [5] N. F. Viemeister, "Psychoacoustic aspects of auditory intensity coding," in *Auditory Function*, eds. G. M. Edelman, W. E. Gall and W. M. Cowan, New York, John Wiley and Sons: 1988.
- [6] B. W. Edwards, "The effect of fast-acting compression on psychoacoustic jnds," *J. Acoust. Soc. Am.*, vol. 103(S), p. 2977.
- [7] B. R. Glasberg, B. C. J. Moore and S. P. Bacon, "Gap detection and masking in hearing-impaired and normal-hearing subjects," *J. Acoust. Soc. Am.*, vol. 81, pp. 1546-1556.