Audio Engineering Society

Convention e-Brief 620

Presented at the 149th Convention
Online, 2020 October 27-30

Investigation of a real-time hearing loss simulation for use in audio production

Angeliki Mourgela[1], Trevor R. Agus[2], and Joshua D. Reiss[1]

[1] *Queen Mary University of London, London*

[2] *Queen's University, Belfast*

Correspondence should be addressed to Angeliki Mourgela (a.mourgela@qmul.ac.uk)

ABSTRACT

We present a perceptually motivated, real-time hearing loss simulation for use in audio production. The implementation builds on a previous simulation, but is now real-time, low latency, and available as a stereo audio effect plug-in with more accurate modelling of hearing loss. It offers the option of isolating and customizing high frequency threshold attenuations on each ear, corresponding to the audiogram information. The simulation also provides the option of incorporating additional suprathreshold effects such as spectral smearing, rapid loudness growth and loss of temporal resolution on audio. The underlying psychoacoustic principles are described, and results are presented to show the simulation's performance.

1 Introduction

Hearing loss simulations have been useful tools in psychoacoustic and audiological research throughout literature [1], since they can provide an easy way of reproducing and assessing the different effects of hearing loss on normal hearing listeners, as well as evaluating the efficiency of different methods for audio enhancement [2]. However, their use in audio production to this day appears to be limited.

This paper investigates the use of a real-time, customizable hearing loss simulation, in the form of a stereo audio plugin effect, for easier incorporation into audio production, as a referencing and analysis tool. The simulation aims to reproduce four perceptual aspects of hearing loss, as described in psychoacoustic research models of auditory perception, by utilising data and existing methods of simulation found in literature. Details on the design and implementation of the four modules comprising the simulation, as well as results showcasing the simulation's performance, are provided in the following sections.

2 Background

Hearing loss is a high prevalence global phenomenon, with approximately 466 million people worldwide suffering from disabling loss, and an estimated increase to 900 million by the year 2050[3]. Causes of hearing loss include genetics, diseases, medication, trauma and deterioration due to ageing, while it is currently estimated that 1.1 billion people of younger age are at risk of developing hearing loss, due to excessive sound level exposure in recreational settings. Daily and multi-hour use of personal listening devices further contributes to this risk [3], [4], [5].

Hearing loss can severely impact the daily life of an individual, causing both functional and emotional difficulties and affecting their overall quality of life, thus making research efforts towards a better understanding of its physical and perceptual characteristics, as well as the development of new and efficient methods for audio enhancement, an essential endeavor for the future [6], [7], [8], [9].

Hearing loss simulations have been used in psychoacoustic and audiological research studies [1], [10],[13], in order to provide researchers with a useful tool when exploring the perceptual aspects of hearing loss. Using normal hearing listeners with simulated hearing loss is an established way of assessing performance and quality perception differences. This way researchers can use performance comparisons between simulated and real hearing losses in order to further assess suprathreshold auditory deficiencies—something otherwise reliant only on physical differences between healthy and impacted listeners—as well as the affected individual’s own experience.

Hearing loss simulations can also reduce the need for recruiting listeners with actual loss, thus facilitating a greater number of studies, by limiting the necessity for potentially difficult-to-access participants. Simulations could also provide a useful audio quality analysis and evaluation tool in audio production and enhancement, especially when assessing audio clarity and intelligibility for audiences with hearing loss, by evaluating the output of the simulation, either using existing objective audio and speech quality metrics (PEAQ, PESQ), or through conducting listening tests [14], [15].

Studies have shown that individuals with hearing loss can experience a degradation in sound clarity which in turn could have a negative impact on speech intelligibility, when stimuli are presented along with a fluctuating masker[16],[17]. This phenomenon is commonly observed in situations such as a speech signal presented along ambient noise or babble, as well as in situations where the acoustic properties of a sound are rapidly changing (e.g. television sound).

This paper aims to investigate the use of hearing loss simulations in audio production, as a tool for audio referencing, analysis and evaluation of complex and constantly changing stimuli. This development builds upon a previous prototype design [18] but with added features, including real-time capability, a more compact and intuitive design with retained customisability features, as well as a stereo configuration. The simulation is exported in a stereo plugin format, therefore making it easier to be incorporated in the stage of audio production. As a mixing tool, this plugin can be inserted either on individual channels, or the stereo master bus of a digital audio workstation, in order to provide the engineer with a quick way of assessing their mix’s quality, when hearing loss related degradations are presented. This way accessibility adjustments are encouraged early on at the stage of production.

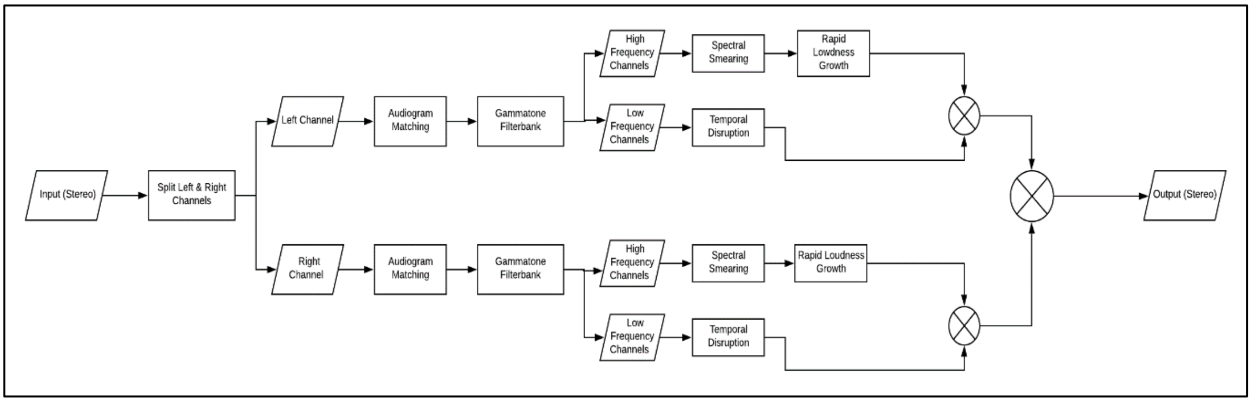

Figure 1: Block diagram of the simulation's processing algorithm

3 Technical Description

The development of this real time simulation encapsulates four different processors: high frequency attenuation, spectral smearing, rapid loudness growth (RGL) and temporal disruption, each replicating a specific perceptual aspect of hearing loss. The diagram in Figure 1 presents the signal flow and processing inside the audio plugin class. The simulation is designed using the audio plugin functionality found in MATLAB's Audio Toolbox and is available in a audio plugin format with MacOS and Windows compatibility. Available at: https://code.soundsoftware.ac.uk/hg/hearing-loss-simulation-plugin.

Processing begins by separating the audio input into a left and right channel, representing the left and right ear. Upon that, the interface offers two options of audiogram attenuation for each of the two channels, corresponding to a mild or a moderate loss. Attenuation is provided at the seven frequencies (250, 500, 1000, 2000, 4000 and 8000 Hz) typically measured in an audiometry test. This stage aims to replicate threshold shifts, often attributed to the loss or compromise of the function of the outer hair cells [19]. To ensure attenuation of frequencies above the audiometric range (>8000 Hz), a low pass filter is also applied with a cut-off frequency of 8000 Hz and a slope of 48 dB. The attenuated audio of each channel is then passed through a Gammatone Filterbank, which separates each of the two channels into 32 equivalent rectangular bandwidth (ERB) spaced bands, ranging between 20 Hz and 16000 Hz. The use of the Gammatone Filterbank at this stage is to ensure that band separation in the plugin corresponds to the separation typically observed in the human cochlea [19]. After band separation, the high and low frequencies of each of the two original channels are further divided before they are sent to the following processors.

High frequencies - The frequency smearing processor performs a multiplication of each of the high frequency bands of the signal, with a low passed white noise. This results in a 'smeared' representation of the frequency content of the sound, which aims to replicate the effects of spectral smearing, usually observed as a result of the widening of the auditory filters in hearing loss, that causes decreased frequency selectivity[12]. The frequency smearing processor offers the user with 3 processing options:

1. **Bypass** - no smearing is applied, and signal is passed unaffected
2. **Low** - a low level of smearing is applied. The level of smearing corresponds to the cut-off frequency of the low-passed white noise-in this case 100 Hz.
3. **High** - a high level of smearing is applied. The level of smearing corresponds to the cut-off frequency of the low-passed white noise-in this case 200 Hz.

The high frequency bands of both ears are then passed through a rapid loudness growth (RLG) processor. This processor aims at replicating the effect of loudness recruitment, a phenomenon resulting from the loss of dynamic range, also observed in individuals with outer hair cell damage [21]. Consequences of RLG include the perception of a more linear and steep growth of loudness with increasing sound level above threshold of hearing, as opposed to the compressive nonlinear growth observed in normal hearing listeners [22]. To achieve this, the processor utilises an upwards expansion approach, where if the signal level goes above a set threshold of hearing and below a set threshold of total recruitment it becomes amplified, whereas if the signal level reaches the threshold of total recruitment the processor returns to an in=out behaviour.

Low frequencies - Parallel to the high frequency bands, the low frequency bands of both ears run through a temporal disruption processor, which aims to recreate loss of temporal resolution. The approach behind this processor is based on observations in literature suggesting that listeners with hearing loss are unable to utilize temporal fine structure (TFS) as well as normal hearing listeners [23], [24]. To achieve the perceptual consequences of reduced access to TFS information, the module applies random phase shifts on the input. More specifically, the signal is analysed through use of a Short Time Fourier Transform (STFT) in MATLAB, using a 512 sample Hann periodic window. The magnitude of the signal is kept intact, while random shifts are applied to the phase of the signal within the range of $[-\pi/2, \pi/2]$. The original magnitude is then recombined with the shifted phase, before the signal gets reconstructed using an Inverse Short Time Fourier Transform (ISTFT). The goal of this processor is to disrupt the periodicity of the signal [25]. Following their separated processing, high and low frequency bands of each channel are re-combined and both channels are brought together to form the stereo output. The plugin also offers bypass options on each processor, as well as a general bypass for each ear. A mute option for each of the two ears is also provided. The graphic user interface of the plugin can be seen in Figure 2.

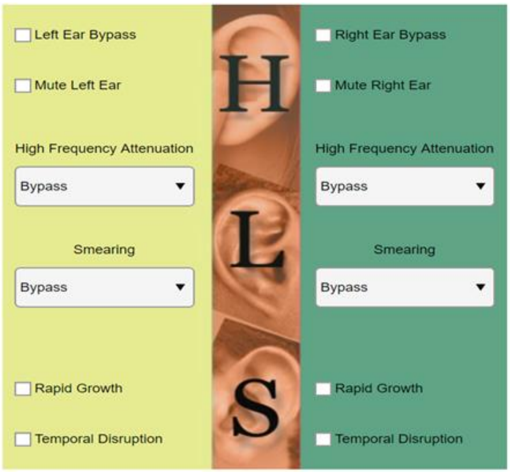

Figure 2: Hearing loss simulation plugin, graphic user interface (GUI)

4 Results

The following section presents results describing the response of the simulation with test signals for each of the plugin's processors. Each processor is tested independently in order to assess its proximity to the desired response.

High Frequency Attenuation - This processor aims to recreate the high frequency attenuation found in a typical audiogram. Figure 3 depicts spectral analysis of white noise passing through the high frequency attenuation processor using a *mild* setting.

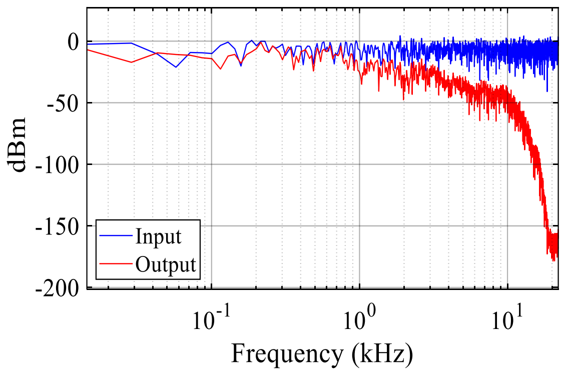

Figure 3: White noise with high frequency attenuation (*mild* setting)

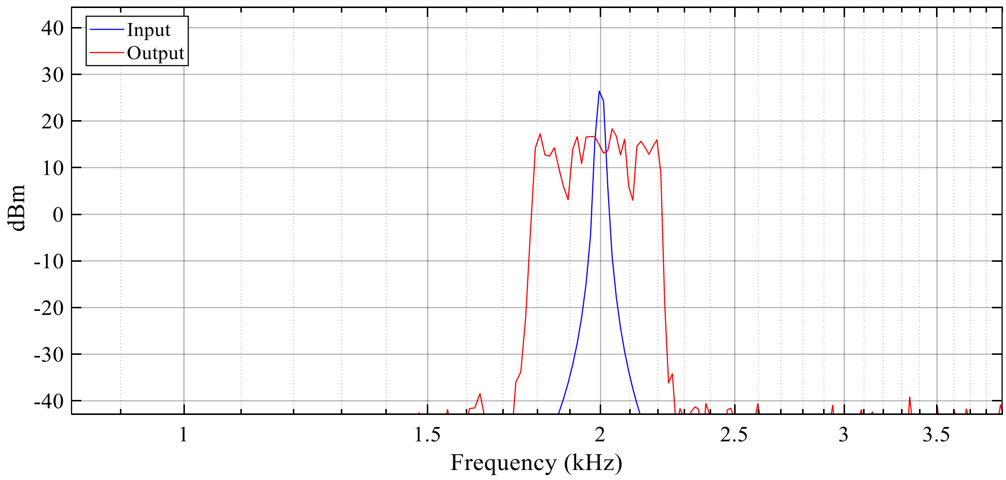

Figure 4: 2kHz tone with smearing (*high* setting)

Spectral Smearing - This processor aims to reproduce the perceptual effect of spectral "smearing", occurring as a consequence of reduced frequency selectivity in the cochlea, due to widening of the auditory filters [26]. The plot in Figure 4 presents the frequency spectrum analysis of a 1 kHz pure tone before and after the *high* smearing setting is applied.

Rapid Loudness Growth (RLG) - This processor aims to recreate the effect of loudness recruitment occurring in the cochlea as a result of outer hair cell damage and restricted or absent function of the active mechanism of the cochlea as described above [9], [22]. The plot in Figure 5 presents the amplitude of a ramp signal plotted over time in logarithmic scale, before and after the RLG processor. An arbitrary threshold of hearing is given here at -55 dB with the attack time set at 3 ms and release time set at 2 ms. For the purpose of focusing on the effect of rapid loudness growth, attenuation below threshold is not presented on this plot.

Loss of Temporal Resolution - The aim of this processor is to disrupt periodicity by introducing random phase shifts, thus making TFS information less available to a listener, corresponding to findings supporting that use of TFS information in listeners with hearing loss is limited [23], [24]. The plots in Figure 6 present the time domain analysis of the application of temporal disruption on a sinewave. As it can be observed there is a fixed delay between the input and the output, which is due to the processing latency of the plugin. A high frequency noise as well as a slight attenuation can also been seen, which are both artefacts of the phase shifting occurring in the TFS disruption processor.

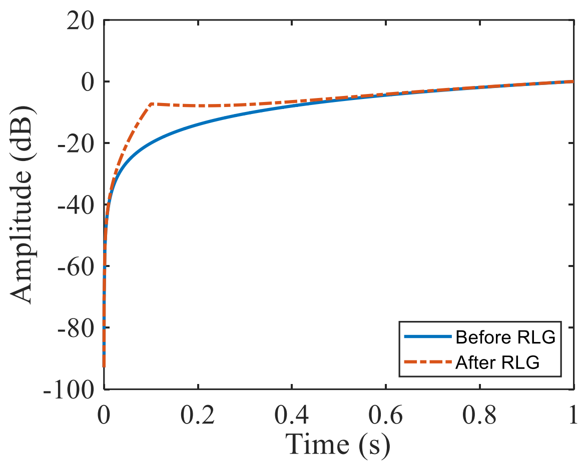

Figure 5: Amplitude of over time plot in logarithmic scale, before and after the RLG processor (ramp signal).

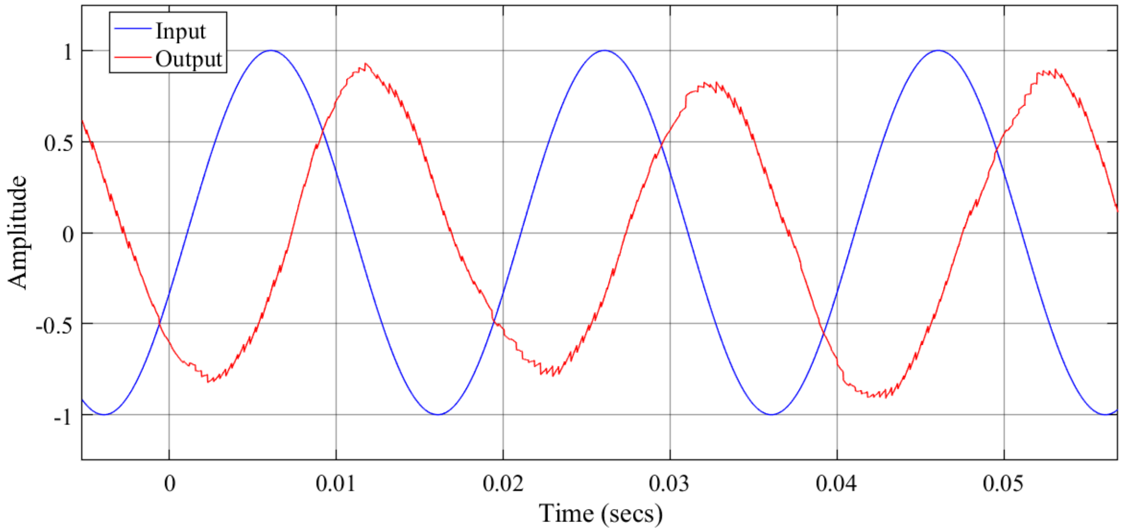

Figure 6: 50 Hz tone with temporal disruption

5 Discussion

The development of a real-time hearing loss simulation was presented, along with results on its performance and response. The results presented in the previous section, demonstrate that perceptual aspects of hearing loss, as documented in current literature, can be approximated and applied on audio signals in real time and with the option of customisation, however presenting certain limitations and considerations.

As with all similar developments, this hearing loss simulation comes with limitations in its performance and effectiveness. Such limitations include a difficulty in reproducing all perceptual aspects of hearing loss accurately, since their documentation in literature mainly relies on performance comparisons on simple audio tasks, between normal listeners and listeners with hearing loss in controlled environments, as well as physical modelling and affected individuals' own descriptions of their losses' characteristics, which make it difficult to establish a reliable reference.

Evaluation of hearing loss simulations can therefore only be made to a limited extend and is performed mainly to ensure that at a minimum there are similar tendencies between the degradation in performance of normal listeners using the simulation and listeners with actual hearing loss.

Another limitation of this development arises from the sacrifices made in the complexity of the digital signal processing (DSP)used in the implementation, in order to ensure good real-time performance, thus resulting on occasionally poorer approximation and accuracy. The use of more complicated and computationally expensive digital signal processing approaches can be found in many hearing loss simulations throughout literature, however most of these models are used to pre-process stimuli offline, which makes considerations such as latency and computational costs irrelevant for their purpose. Additional limitations include the occurrence of DSP-related unwanted artefacts, resulting from processing operations affecting multiple aspects of the signal.

An important consideration regarding the approach presented in this paper lies within its intended use, which is to promote awareness of the quality and accessibility of audio content for all, while facilitating adjustments at the stage of audio production. It is therefore designed to be used as a tool for referencing and analysis, rather than as an attempt to accurately model hearing loss. Future directions of this project include conducting perceptual evaluation listening tests to assess its performance using real listeners, as well as an investigation of its use in practice, in order to test the functionality and capability of the simulation to assist the development of effective audio enhancement methods for listeners with hearing loss.

6 Acknowledgements

This work is supported by the EPSRC and the BBC.

References

[1] J. P. Gagne and N. P. Erber, “Simulation of sensorineural hearing impairment,” *Ear Hear.*, 1987, doi: 10.1097/00003446-198708000-00008.

[2] Isaac John Graf, “Simulation of the Effects of Sensorineural Hearing Loss,” Massachusetts Institute of Technology, 1997.

[3] World Health Organisation, ‘Deafness and hearing loss’, March 1, 2020, https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

[4] J. H. Peng, Z. Z. Tao, and Z. W. Huang, “Risk of damage to hearing from personal listening devices in young adults,” *J. Otolaryngol.*, 2007, doi: 10.2310/7070.2007.0032.

[5] A. H. Sulaiman, R. Husain, and K. Seluakumaran, “Evaluation of early hearing damage in personal listening device users using extended high-frequency audiometry and otoacoustic emissions,” *Eur. Arch. Oto-Rhino-Laryngology*, 2014, doi: 10.1007/s00405-013-2612-z.

[6] B. Shirley and R. Oldfield, “Clean audio for TV broadcast: An object-based approach for hearing-impaired viewers,” *AES J. Audio Eng. Soc.*, 2015, doi: 10.17743/jaes.2015.0017.

[7] B. Shirley, L. Ward, and E. T. Chourdakis, “Personalization of object-based audio for accessibility using narrative importance,” in *CEUR Workshop Proceedings*, 2019.

[8] T. R. A. and C. Corrigan, “Adapting Audio Mixes for Hearing Impairements,” 2017.

[9] L. A. Ward and B. G. Shirley, “Personalization in object-based audio for accessibility: A review of advancements for hearing impaired listeners,” *AES J. Audio Eng. Soc.*, 2019, doi: 10.17743/jaes.2019.0021.

[10] R. P. Derleth, T. Dau, and B. Kollmeier, “Modeling temporal and compressive properties of the normal and impaired auditory system,” *Hear. Res.*, 2001, doi: 10.1016/S0378-5955(01)00322-7.

[11] M. Cuevas-Rodriguez *et al.*, “The 3D Tune-In Toolkit - 3D audio spatialiser, hearing loss and hearing aid simulations,” in *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments, SIVE 2018*, 2018, doi: 10.1109/SIVE.2018.8577076.

[12] J. G. Desloge, C. M. Reed, L. D. Braida, Z. D. Perez, and L. A. Delhorne, “Auditory-Filter Characteristics for Listeners With Real and Simulated Hearing Impairment,” *Trends Amplif.*, 2012, doi: 10.1177/1084713812445510.

[13] R. K. Muralimanohar, J. M. Kates, and K. H. Arehart, “Using envelope modulation to explain speech intelligibility in the presence of a single reflection,” *J. Acoust. Soc. Am.*, 2017, doi: 10.1121/1.4983630.

[14] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, “Perceptual evaluation of speech quality (PESQ) - A new method for speech quality assessment of telephone networks and codecs,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2001, doi: 10.1109/icassp.2001.941023.

[15] T. Thiede *et al.*, “PEAQ - the ITU standard for objective measurement of perceived audio quality,” *AES J. Audio Eng. Soc.*, 2000.

[16] P. E. Souza and C. W. Turner, “Masking of speech in young and elderly listeners with hearing loss,” *J. Speech Hear. Res.*, 1994, doi: 10.1044/jshr.3703.655.

[17] J. M. Festen and R. Plomp, “Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing,” *J. Acoust. Soc. Am.*, 1990, doi: 10.1121/1.400247.

[18] A. Mourgela, T. Agus, and J. D. Reiss, “Perceptually motivated hearing loss simulation for audio mixing reference,” in *147th Audio Engineering Society International Convention 2019*, 2019.

[19] W. C. Stebbins, J. E. Hawkins, L. G. Johnsson, and D. B. Moody, “Hearing thresholds with outer and inner hair cell loss,” *Am. J. Otolaryngol. Neck Med. Surg.*, 1979, doi: 10.1016/S0196-0709(79)80004-6.

[20] B. R. Glasberg and B. C. J. Moore, “Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments,” *J. Acoust. Soc. Am.*, 1986, doi: 10.1121/1.393374.

[21] M. K. Pichora-Fuller and E. MacDonald, “Sensory aging: Hearing,” in *The Curated Reference Collection in Neuroscience and Biobehavioral Psychology*, 2016.

[22] A. J. Oxenham and S. P. Bacon, “Cochlear compression: Perceptual measures and implications for normal and impaired hearing,” *Ear and Hearing*. 2003, doi: 10.1097/01.AUD.0000090470.73934.78.

[23] C. Lorenzi, G. Gilbert, H. Carn, S. Garnier, and B. C. J. Moore, “Speech perception problems of the hearing impaired reflect inability to use temporal fine structure,” *Proc. Natl. Acad. Sci. U. S. A.*, 2006, doi: 10.1073/pnas.0607364103.

[24] K. Hopkins and B. C. J. Moore, “Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information,” *J. Acoust. Soc. Am.*, 2007, doi: 10.1121/1.2749457.

[25] M. S. Sommers and L. E. Humes, “Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listenersA,” *J. Acoust. Soc. Am.*, 1993, doi: 10.1121/1.405810.