We employ computational models of loudness and pitch perception to better understand the impact of sensorineural hearing loss on music perception, with the aim of guiding technology development for hearing-impaired listeners. Traditionally, hearing aid development has been geared towards improving speech intelligibility and has largely failed to provide adequate restoration of music to those with hearing loss. One difficulty with trying to improve music perception in impaired listeners is the absence of a good quantitative measure of music reception, analogous to speech reception measures like word-recognition rate. Psychoacoustic models for loudness and pitch allow us to gauge quantitative parameters relevant to music perception and make predictions about the type of deficits listeners face. We examine the impact of hearing loss to predicted measures of loudness, specific loudness, pitch and consonance and make suggestions on possible methods for restoration.
INTRODUCTION

Traditionally, hearing aid development has been geared towards improving speech intelligibility and has largely failed to provide adequate restoration of music to those with hearing loss. One difficulty with trying to improve music perception in listeners with hearing loss is the absence of quantitative measures of music reception, analogous to speech reception measures like word-recognition rate. However, psychoacoustic models for loudness and pitch allow us to gauge quantitative parameters relevant to music perception and make predictions about the types of deficits listeners face. In this study, we employ computational models of loudness and pitch perception to better understand the impact of sensorineural hearing loss on music perception, with the aim of guiding technology development for listeners with hearing loss. Specifically, we examine the impact of hearing loss on predicted measures of loudness, specific loudness, pitch and consonance and make suggestions on possible methods for restoration of these attributes of music.

METHODS

We generated model predictions of loudness and pitch salience for a number of stimuli over a range of hearing loss types and standard amplification strategies.

Models

Loudness Model

We used an established model for loudness of time-varying sounds (Glasberg and Moore, 2002) to predict the perceived overall loudness and frequency-specific loudness (“specific loudness”) for listeners with hearing loss and listeners with normal hearing. This model has been used to predict a wide variety of loudness phenomena for both normal listeners and listeners with hearing loss (Moore and Glasberg, 2004). In this model, “instantaneous” specific loudness is estimated frame by (1 ms) frame, based on the predicted instantaneous auditory excitation pattern. Specific loudness is integrated (summed) across frequency to compute and estimate the instantaneous overall loudness. “Instantaneous loudness” is assumed to be perceptually inaccessible, and the instantaneous values are smoothed asymmetrically to obtain short-term loudness predictions. Single-pole smoothing time constants of approximately 22 ms attack – loudness increasing -- and 50 ms release – loudness decreasing – are used to perform the smoothing (predicted perception of loudness increases more quickly than it decreases). The short-term loudness is further smoothed (time constants 100 ms attack and 200 ms release) to predict the perception of long-term loudness.

Pitch Model

For pitch salience, we used a recent model (Bidelman and Heinz, 2011) that has been used to predict the relative pitch salience for diatonic musical intervals. Additionally, the model predicts the reduction in pitch salience contrast seen in listeners with hearing loss (Tufts et al., 2005). The model is a temporal model for pitch, based on auditory nerve fiber inter-spike intervals. It is comprised of an auditory nerve model (Zilany and Bruce, 2006) followed by a periodic sieve template (Larsen et al., 2008). Inter-spike intervals are calculated using the autocorrelation function (ACF) on peri-stimulus time histograms of auditory-nerve responses (spikes) and then passed through the periodic sieve template to determine the relative proportion of intervals at a particular pitch (F0) period. Prior to the periodic sieve template, the ACF is weighted to provide greater emphasis on shorter intervals, a likely consequence of neural processing to obtain the ACF (Bernstein and Oxenham, 2005). Pitch salience of a particular F0 is taken to be the relative proportion of inter-spike intervals at that F0.

Stimuli

Loudness analyses were performed using four 15-second excerpts of Western music varying in style (classical, country, pop and jazz), instrumentation and dynamics. Stimuli for the pitch modeling analyses included harmonic complex tones and dyads as described in (Bidelman and Heinz, 2011) and iterated ripple noises as described in (Leek and Summers, 2001).
Procedure

Analyses were based on nine audiograms: three classified as flat losses (similar loss at low and high frequencies); three classified as gradually sloping (loss gradually increasing from low to high frequencies); and three classified as steeply sloping (loss increases rapidly from low to high frequencies). Two types of amplification strategies were fit for nine different audiograms using in receiver-in-the-canal (RIC) style hearing aids; we used both a linear amplification strategy and a compressive amplification strategy intended for music listening. Hearing aids were setup on a KEMAR manikin and monaural recordings were made through the various amplification treatments, along with unaided conditions. The recordings were fed through the loudness and pitch models using the appropriate hearing-loss parameters for each configuration and analyses were performed on the outputs of the models as described below.

Loudness Model Analyses

We evaluated overall loudness and a number of statistics of specific loudness. The specific loudness statistics were derived through definitions of spectral quality attributes from recording engineering terminology, such as loudness balance and fullness. These were then applied to the specific loudness profiles.

Pitch Model Analyses

Because of the novelty of the pitch salience model we ran some basic stimuli through it to examine the effect of hearing loss and stimulus parameters on model predictions. In addition we examined the predicted pitch salience of complex-tone dyads across an octave range of musical intervals and of iterated ripple noises designed to elicit different degrees of pitch salience.

RESULTS

Loudness

Overall Loudness

Results show, as expected, that all hearing loss types reduce the overall predicted loudness compared to normal-hearing listeners. Results from the amplified conditions show that different amplification strategies boost the overall loudness but not equally. The predicted loudness range of the linear amplification treatment is much greater than normal-hearing loudness: the loud portions of excerpts are much louder than normal and the soft parts are softer than normal. The compressive strategy is much better at restoring the overall loudness and the range of loudness closer to that of normal-hearing listeners.

Loudness Balance and Fullness

Loudness balance, i.e., the balance of specific loudness across the frequency range, is affected differently by the different hearing loss types compared to normal-hearing loudness balance. The spectral shapes of hearing loss types are reflected in the general shapes of specific loudness profiles, however loudness balance changes with level – much more than for normal-hearing listeners. The aided conditions show that compressive amplification better restores overall loudness balance and the consistency of loudness balance across stimulus level.

Predicted fullness is defined as the ratio of specific loudness in the 250-500 Hz range to that in the 250-2000 Hz range. Results show that for normal-hearing listeners, predicted fullness varies over time but is quite stable across level. For listeners with hearing loss, predicted fullness is comparatively more variable over time and level. Both amplification strategies help stabilize the variability of fullness to some degree, but they do not completely restore to the normal range.
Pitch Salience

Harmonic Tone Complexes

Model predictions for harmonic complex tones (6 harmonics, flat spectrum) across a range of levels show that predicted pitch salience is reduced overall for listeners with hearing loss but grows more rapidly as a function of level and is greater than normal-hearing pitch salience at high levels. This counter-intuitive effect at high levels depends on the fundamental frequency and relative phase of the harmonics and appears to stem from a heightened sensitivity to temporal envelope modulations in high characteristic-frequency auditory nerve fibers. Due to this effect, we restricted our use of the model to moderate stimulus levels.

Harmonic Tone Complex Dyads

Model predictions show, as in (Bidelman and Heinz, 2011), that hearing loss reduces the relative differences in pitch salience across musical intervals. Our results also show that amplification increases the pitch salience to normal levels for many hearing-loss profiles and there appears to be no general advantage of one amplification strategy over the other.

Iterated Ripple Noises

Leek and Summers (2001) show a clear and significant difference in pitch strength of iterated ripple noises for normal-hearing listeners compared to listeners with hearing loss. Our model predictions of the same stimuli show very weak pitch salience overall, even at the extreme stimulus parameter conditions designed to provide the most salient pitch. In addition, there was little difference between model predictions for listeners with hearing loss and normal-hearing listeners.

DISCUSSION

We used a well-established loudness model that has been used to predict a wide variety of loudness-based phenomena including loudness recruitment in listeners with hearing loss. Our interpretation and particular analyses of specific-loudness to derive predicted qualities such as fullness are, to some degree unverified but have promise based on our understanding of the loudness model. Our results could be made more robust with corresponding psychoacoustic data to match the specific loudness qualities. In addition, we used default loudness model parameters for the ratio of inner to outer hair cell loss, which, if derived through perceptual measures could lead to more precise model predictions for individual hearing losses. Finally, it is important to remember that loudness-based hearing aid fittings are spectra specific and may not correspond to optimal loudness restoration for all spectral levels and shapes. Nevertheless, even with these caveats in mind, our results suggest some general guiding principals for restoring loudness and specific loudness qualities in listeners with hearing loss.

The pitch model we employed was a natural extension of previous work but our results suggest that it needs further development. In particular the sensitivity of pitch salience predictions to the temporal envelope must be addressed. In addition, we need more data on pitch salience of listeners with hearing loss.

As we develop accurate models for music-specific perceptual qualities, we can more effectively predict the effect of specific amplification strategies and signal-processing techniques intended to improve music listening through hearing aids. The model predictions can be used to indicate if these perceptual attributes have been restored to within normal operating ranges.

REFERENCES


