Individual Differences Reveal Correlates of Hidden Hearing Deficits

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1523/jneurosci.3915-14.2015">http://dx.doi.org/10.1523/jneurosci.3915-14.2015</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Society for Neuroscience</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sat Feb 16 02:29:48 EST 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/98035">http://hdl.handle.net/1721.1/98035</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Individual Differences Reveal Correlates of Hidden Hearing Deficits

Hari M. Bharadwaj,1,2 Salwa Masud,1,2 Golbarg Mehraei,1,3 Sarah Verhulst,1,4 and Barbara G. Shinn-Cunningham1,2

1Center for Computational Neuroscience and Neural Technology and 2Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, 3Program in Speech and Hearing Biosciences and Technology, Harvard University–Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and 4Cluster of Excellence Hearing4all and Medizinische Physik, Department of Medical Physics and Acoustics, Oldenburg University, D-26111 Oldenburg, Germany

Clinical audiometry has long focused on determining the detection thresholds for pure tones, which depend on intact cochlear mechanics and hair cell function. Yet many listeners with normal hearing thresholds complain of communication difficulties, and the causes for such problems are not well understood. Here, we explore whether normal-hearing listeners exhibit such suprathreshold deficits, affecting the fidelity with which subcortical areas encode the temporal structure of clearly audible sound. Using an array of measures, we evaluated a cohort of young adults with thresholds in the normal range to assess both cochlear mechanical function and temporal coding of suprathreshold sounds. Listeners differed widely in both electrophysiological and behavioral measures of temporal coding fidelity. These measures correlated significantly with each other. Conversely, these differences were unrelated to the modest variation in otoacoustic emissions, cochlear tuning, or the residual differences in hearing threshold present in our cohort. Electroencephalography revealed that listeners with poor subcortical encoding had poor cortical sensitivity to changes in interaural time differences, which are critical for localizing sound sources and analyzing complex scenes. These listeners also performed poorly when asked to direct selective attention to one of two competing speech streams, a task that mimics the challenges of many everyday listening environments. Together with previous animal and computational models, our results suggest that hidden hearing deficits, likely originating at the level of the cochlear nerve, are part of “normal hearing.”

Key words: central auditory processing disorders; cochlear neuropathy; envelope-following response; frequency-following response; hidden hearing loss; temporal coding

Introduction

Clinical practice characterizes hearing loss through pure-tone threshold audiometry. However, a number of normal-hearing (NH) listeners complain of difficulty communicating, especially when there are competing sounds, noise, and room reverberation. When such listeners seek clinical treatment, they are labeled as suffering from obscure auditory dysfunction (Saunders and Haggard, 1992), King–Kopetzky syndrome (Zhao and Stephens, 1996), or, more generally, (central) auditory processing disorders (Chermak and Musiek, 1997). The nonspecific nature of these labels underscores how little we understand about the underlying mechanisms.

As many as 5–15% of adult patients seeking audiological help turn out to have normal thresholds (Cooper and Gates, 1991; Kumar et al., 2007; Hind et al., 2011). Importantly, the problem may be even more common than such figures suggest; for instance, some NH listeners, particularly older participants, perform poorly on behavioral tasks that rely on precise temporal cues in suprathreshold sound (Strelcky and Dau, 2009; Grose and Mamo, 2010; Ruggles et al., 2012). Auditory temporal perception degrades with age, independently of hearing loss and throughout adulthood (Snell and Frisina, 2000; Snell et al., 2002; Fitzgibbons and Gordon-Salant, 2010; Ruggles et al., 2012). Many listeners who can communicate effectively in quiet settings are nonetheless frustrated in complex environments such as cocktail parties, noisy restaurants, and busy streets (Dubno, 1984; Gatehouse and Noble, 2004). These findings suggest that a significant portion of listeners with normal or near-normal hearing thresholds exhibit temporal processing deficits.

Previous animal work revealed that noise exposure causing temporary threshold shifts, previously thought to be innocuous, can produce a loss of as many as 50% of the synapses and cochlear-nerve terminals innervating the inner hair cells, described herein as cochlear neuropathy (Kujawa and Liberman, 2009; Lin et al., 2011b). Aging also appears to reduce the effective auditory nerve fiber (ANF) population, initially through a loss of synapses with a subsequent degeneration of cell bodies (Makary...
et al., 2011; Sergeyenko et al., 2013). Furthermore, significant levels of deafferentation do not appear to affect detection thresholds (Schuknecht and Woelwer, 1955; Lobarinas et al., 2013); however, in the healthy auditory system, convergence of nerve fibers is important for precise coding of both temporal fine-structure and envelope cues in suprathreshold sounds (Joris et al., 1994; Lopez-Poveda and Barrios, 2013). Given the importance of temporal information for speech perception (Zeng et al., 2005), source localization (Blauert, 1997), grouping of acoustic elements into perceptual objects (Darwin, 1997; Shamma et al., 2011), and release from masking (Moore, 2008; Christiansen et al., 2013), degraded temporal coding is likely to interfere with suprathreshold perceptual ability—a form of “hidden hearing loss” (Bharadwaj et al., 2014; Plack et al., 2014).

Using a combination of electrophysiological, otoacoustic, and perceptual measures, the current study exploits individual differences in a cohort of NH listeners to characterize the relationships between peripheral function, neural temporal coding, and suprathreshold hearing performance with complex stimuli. We find that suprathreshold temporal coding fidelity varies in listeners with no known hearing complaints in a pattern consistent with cochlear neuropathy.

Materials and Methods
A primary goal of the study was to determine whether the pattern of individual differences in suprathreshold temporal coding in a cohort of young NH adult listeners is consistent with cochlear neuropathy. We explicitly evaluated peripheral cochlear mechanics and hair cell function to disambiguate where differences arise. Because noise-induced temporary and permanent threshold shifts often present as notches around 4 kHz (Pierson et al., 1994; Yost, 2007), our stimuli and testing focused on this frequency region.

Methods and measures were designed to imitate the challenges facing listeners in everyday, noisy environments in controlled laboratory settings. In contrast to listening to a single speech source in quiet (which is sparse in both time and frequency), acoustic scenes with multiple talkers, noise, and room reverberation typically have the following:

1. High overall sound levels. A larger percentage of the overall population of ANFs will respond, including high-threshold, low–spontaneous rate (SR) fibers that would contribute less at low sound intensities.

2. A relatively dense pattern of cochlear excitation along the tonotopic axis. “Off-frequency” ANF channels, remote from the dominant stimulus-evoked excitation locations, cannot encode target information. This is in contrast to listening in quiet where off-frequency high-SR fibers would respond to target modulations (Fig. 1).

3. Relatively weak signal modulation. In addition to directly masking the target modulations, energy from competing sources will tend to fill in temporal dips in signal energy, requiring robust encoding of small modulations of high-intensity sounds for target information to be discernible (Fig. 1).

These observations suggest that at the level of the auditory nerve, degradations in temporal coding fidelity may be evident when listening in noisy real-world conditions even though they may not impair perception in conditions like those often tested in the laboratory. We therefore employed high sound levels, off-frequency maskers, and stimuli with shallow modulation depths in our battery of measures (Table 1). Given that neuropathy from both noise exposure and aging preferentially (though by no means exclusively) impacts low-SR fibers (Fig. 2), we hypothesized that such stimuli would increase the likelihood of exposing suprathreshold temporal coding deficits resulting from neuropathy (Bharadwaj et al., 2014). We measured peripheral/cochlear processing of sounds using distortion-product otoacoustic emissions (DPOAEs; Johnson et al., 2006), psychophysical tuning curves (Oxenham and Shera 2003), and audiometric detection thresholds. We measured early neural coding of temporal information using subcortical envelope-following responses (EFRs) (Bharadwaj and Shinn-Cunningham, 2014; Bharadwaj et al., 2014). We then compared these measures to perceptual estimates of temporal sensitivity, to electroencephalography (EEG)-based measures of cortical representation of acoustic spatial features (Salinen et al., 2010), and to selective attention performance in a complex speech-on-speech masking task. Finally, we compared perceptual measures of temporal sensitivity to self-reported noise-exposure history.

All measures were obtained with the participants seated in an acoustically and electrically shielded booth (single-walled Eckel C-14 booth). For passive measures (DPOAEs and EFRs), participants watched a silent, captioned movie of their choice, ignoring the acoustic stimuli. A personal desktop computer controlled all aspects of the experiment, including triggering sound delivery and storing data. Special-purpose sound-control hardware (System 3 real-time signal processing systems, including analog-to-digital conversion and amplification; Tucker Davis Technologies) presented sound through insert phones (ER-1; Etymotic; for all measures except otoacoustic emissions, which used ER-10c) coupled to foam ear tips. All sounds were digitized at a sampling rate of 48,828 Hz. For all active behavioral experiments, subjects responded by button presses. Feedback was presented after each trial.

Participants
Twenty six (13 female) subjects, aged 21–39 years, were recruited from the Boston University community. All subjects had pure-tone hearing thresholds better than 15 dB hearing level in both ears at octave frequencies between 250 Hz and 8 kHz and rated their own hearing as normal and similar to others in their age group. In addition, two subjects (aged 27 and 29) with a “notch”-type hearing loss at 4 kHz were recruited to evaluate the sensitivity of the measures used to characterize cochlear mechanical function. Subjects provided informed consent in accordance with protocols established at Boston University. Although the goal was to obtain each of the measures (Table 1) on every subject, subject availability precluded some subjects from completing some measures (see Fig. 4, Fig. 6).

Correlates of cochlear mechanical function
Psychophysical tuning curves. An important aspect of healthy basilar membrane processing is good frequency selectivity. Psychophysical tuning curves were measured for a fixed, low-intensity probe tone (sensation level (SL) of 10 dB) using the notched-noise method (Patterson, 1976) in a forward-masking paradigm (Oxenham and Shera, 2003). Stimulus parameters and testing procedures were identical to those used by Oxenham and Shera (2003). Filter equivalent rectangular bandwidths (ERBs) were estimated by fitting a roex(p,w,t) function (see Oxenham and Shera, 2003).

DPOAE input–output curves. To obtain an objective correlate of cochlear mechanical nonlinearity, DPOAE growth functions were measured as a function of level of the f2 primary tone (f2 = 4 kHz). The frequency and level of the f1 tone were varied according to the formula provided by Johnson et al. (2006) to maximize the level of the DPOAE for each level of the f2 tone. The DPOAE level was estimated at the distortion frequency of 2f1–f2.

Behavioral measures of temporal coding
Amplitude modulation (AM) detection thresholds were obtained using broadband noise stimuli (Viemeister, 1979). The target was a 500 Hz band of noise centered at 4 kHz and modulated at 19 Hz. Unmodulated bands of noise above (up to 10 kHz) and below (down to 20 Hz) served as maskers to reduce off-frequency cues. The target to notched-noise ratio was 10 dB, and the overall level was 80 dB SPL [broadband root mean square (RMS)]. Reference signals were statistically identical but without any amplitude modulation. The stimuli were presented diotically in a two-alternative, forced-choice paradigm (i.e., one interval with the reference signal and another with the target signal) for different modulation depths sampled randomly between 0 and 100%. By using noise stimuli, spectral cues for AM detection were eliminated. Threshold depths for detection were determined using a Bayesian approach, fitting the parameters of a sigmoidal psychometric function to the responses from 1500
trials and calculating the posterior mean threshold using a Monte Carlo Markov chain (MCMC) sampling procedure from the posterior density (Kuss et al., 2005). As in Kuss et al. (2005), a normal and a log-normal prior were used for the location (threshold) and the slope parameters, respectively. The chance level was fixed at 50%, and a prior was used for the lapse rate parameter.

To obtain a binaural measure of temporal coding, we used an envelope interaural time difference (ITD) discrimination task. A half-wave rectified sinusoid at 40 Hz was used to modulate a 4 kHz carrier (called a “transposed” tone; Bernstein and Trahiotis, 2002). The envelope rate of 40 Hz was chosen because we acquired EEG data simultaneously with the task; this rate is known to elicit strong cortical source auditory steady-state responses (ASSRs) in EEG (Picton et al., 2003). The carrier phase was identical in the two ears; the ITD was applied only to the 40 Hz envelope. The stimulus level was set at 75 dB SPL. Off-frequency notched-noise (notch width of 800 Hz) maskers, realized independently in each trial and uncorrelated across the two ears, were presented at a signal-to-noise ratio

<table>
<thead>
<tr>
<th>Measure</th>
<th>Stimulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audiograms</td>
<td>250 Hz to 8 kHz pure tones</td>
</tr>
<tr>
<td>Psychophysical tuning curves</td>
<td>4 kHz tones at 10 dB SL in notched noise</td>
</tr>
<tr>
<td>DPOAE growth curves</td>
<td>4 kHz “f2” frequency at 16 levels</td>
</tr>
<tr>
<td>Modulation detection</td>
<td>Broadband noise with a 500 Hz band around 4 kHz modulated at 19 Hz (80 dB SPL)</td>
</tr>
<tr>
<td>Envelope-ITD discrimination</td>
<td>4 kHz AM tones in notched noise modulated at 40 Hz (75 dB SPL)</td>
</tr>
<tr>
<td>Envelope-following responses</td>
<td>4 kHz AM tones in notched noise modulated at 100 Hz at a fixed level (75 dB SPL) with varying modulation depths (Fig. 5A)</td>
</tr>
<tr>
<td>Cortical EEG</td>
<td>Same as for envelope ITD detection (Fig. 7A)</td>
</tr>
<tr>
<td>“Spatial” attention</td>
<td>Spoken monotonized digit sequences with varying ITDs plus additive 1/f noise (interaurally uncorrelated)</td>
</tr>
</tbody>
</table>
Electrophysiological measures of temporal coding

EFRs were measured in response to 400-ms-long bursts of 100 Hz transposed tones with a carrier frequency of 4 kHz at a level of 75 dB SPL and for varying modulation depths \((m)\). EFRs were obtained using a 32-channel EEG system (Biosemi Active II system) using 1000 presentations of each stimulus with half in each polarity. The interstimulus interval varied uniformly between 550 and 650 ms. The EFR power was estimated in the frequency domain using a complex principal component analysis approach to combine across channels. This approach combines measurements across multiple channels, optimally adjusting for phase disparities in the signal across channels; the approach has been shown to improve measurements across multiple channels, optimally adjusting for phase disparities of subjects whose thresholds appeared to exceed 800 \(\mu\)s, additional trials were included with ITDs up to 1200 \(\mu\)s.

ITD-based selective attention task

To evaluate the contributions of individual differences in cochlear mechanical function and suprathreshold temporal coding to differences in listening performance in a complex task, a “spatial” attention task was used. Spoken digits recorded in-house in the voice of a female speaker were monotonized to 184 Hz (close to the natural pitch of the voice). To simulate some of the constraints imposed by realistic maskers, uncorrelated pink noise (1/f spectrum; Fig. 1A) was added to each ear. The signals were spatialized by imposing simple ITDs. Each trial consisted of two simultaneous sequences of three spoken digits each, with the streams differentiated only by their ITDs, which were symmetrical around the midline (see Fig. 7A). The 40 Hz cortical ASSR was also quantified.

Results

Correlates of peripheral processing

Figure 3A shows the DPOAE growth function for subjects in our NH cohort (black), as well as for two subjects with hearing loss (red lines). The standard error of the mean over the NH cohort is also shown for reference. The results are comparable to published DPOAE input–output curves for NH listeners (Johnson et al.,
2006). The relatively large attenuation in the DPOAE for the two subjects with hearing loss testifies to the sensitivity of the measure to dysfunction in cochlear mechanics. We defined the DPOAE threshold for each subject as the lowest presented level of the f2 primary tone at which the DPOAE level exceeded 10 dB SPL. These DPOAE thresholds were then compared with behavioral audiometric thresholds and with the behavioral measures of suprathreshold temporal perception. Although all of the listeners in our NH cohort have thresholds within the normal range (by design), there are modest residual differences in the thresholds across listeners (see horizontal distribution of data in Fig. 3C). Psychophysical filter widths also were comparable to those of NH listeners in the published literature (Oxenham and Shera, 2003), with a mean ERB of 249 Hz and a standard error of 24 Hz (Fig. 3B). Importantly, the three measures of cochlear amplifier function correlated with each other (ERBs vs pure-tone thresholds, $r = 0.51, p = 0.011$; DPOAE vs pure-tone thresholds, $r = 0.42, p = 0.04$), suggesting that the variance in these measures across subjects is not dominated by measurement noise, but rather reflects individual differences in cochlear function among the listeners in the NH cohort.

**Behavioral correlates of temporal sensitivity**

Similar to previous reports of large individual differences in temporal coding in NH listeners (Ruggles et al., 2011), we found large individual differences in AM thresholds across our NH cohort (see vertical distribution of data in Fig. 3C). The measured thresholds are higher than AM thresholds for noise stimuli reported in the literature (Viemeister, 1979); this discrepancy is likely due to the fact that our narrowband modulation target signal was presented with unmodulated off-frequency bands of masking noise at a high spectral level. Crucially, these differences in AM thresholds appear to be unrelated to the small individual differences in pure-tone thresholds at 4 kHz (Fig. 3C).

Concomitantly, we also found large individual differences in binaural measures of temporal coding. As shown in Figure 4, envelope-ITD sensitivity for a carrier frequency of 4 kHz varies over a wide range across listeners, from 6 to 19 dB relative to 100 μs (i.e., from ~200 to 900 μs). This range is greater and the absolute threshold values higher than typical past reports (Bernstein and Trahiotis, 2002); however, in contrast to most past studies, we presented our stimuli along with relatively high intensity, interaurally uncorrelated off-frequency noise maskers (Bernstein and Trahiotis, 2008). Across listeners, neither mea-
sure of temporal coding fidelity correlated with audiometric threshold measurements; however, the two measures were significantly correlated with each other ($r = 0.78, p = 6.8e^{-6}$; Fig. 4).

Similarly, psychophysical curve bandwidths (ERBs) did not correlate with either AM thresholds ($r = 0.11, p > 0.05$) or envelope-ITD thresholds ($r = 0.21, p > 0.05$) at the 5% false-alarm-rate level.

Subcortical neural correlates of temporal coding

To assess temporal coding precision with a physiological measure, EFR magnitudes were extracted using frequency-domain principal component analysis (Bharadwaj and Shinn-Cunningham, 2014). Figure 5B shows the EFRs obtained for the four deepest modulation depths tested. On average, as the modulation depth decreases, the EFR magnitude decreases. Figure 5B also shows traces for a representative sampling of the individual NH subjects (the first 10 enrolled in the study, in chronological order). Whereas some individuals showed a gradual decrease in the EFR magnitude as modulation depth decreases, others showed a more precipitous decrease leading to weak EFR responses at shallow modulation depths. For each subject, we fit a straight line to the EFR magnitudes obtained with the three largest modulation depths (0, −4, and −8 dB) and considered the slope relating EFR magnitude to modulation depth. As shown in Figure 5B, slopes varied from close to 0 dB EFR change per 4 dB of modulation to 12 dB EFR change per 4 dB of modulation. These slopes correlate significantly with behavioral AM thresholds (Fig. 6A). Subjects whose EFR magnitude dropped precipitously with a decrease in input modulation had poorer AM sensitivity (high AM thresholds). For comparison, we also considered the absolute EFR magnitude at a shallow modulation depth (Fig. 6B). These absolute magnitudes also correlated significantly with behavioral AM thresholds (e.g., EFR at −4 dB vs AM thresholds, $r = 0.53, p = 0.008$); however, the slope of EFR strength versus stimulus modulation depth exhibited a stronger correlation with the perceptual thresholds ($r = 0.68, p = 2.5e^{-4}$). To compare the correlations, we used the T2 test, which has been shown to perform well when comparing dependent correlations (Steiger, 1980). Based on this test, the slope metric was a significantly stronger correlate of AM thresholds than the EFR at any single modulation depth ($p = 0.03$). The EFR slope metric also correlated strongly with the binaural measure of temporal coding (envelope-ITD thresholds, $r = 0.54, p = 0.006$; Fig. 6C). That the EFR slope metric is a strong correlate of supraphreshold temporal sensitivity conforms to the model predictions of simulated low-SR ANF neuropathy (Bharadwaj et al., 2014). These correlations remained strong when other potential covariates were also included as predictors; specifically, a multiple regression model for the AM thresholds revealed that only the EFR slope was a significant predictor ($F_{(1,18)} = 14.3, p = 0.0013$), whereas age ($F_{(1,18)} = 0.007, p = 0.95$), sex ($F_{(1,18)} = 0.08, p = 0.77$), and audiometric thresholds ($F_{(1,18)} = 0.5, p = 0.49$) were not. Note, however, that our age range was narrow, with the distribution skewed toward the younger side (only three subjects were older than 30).

Cortical EEG correlates of binaural temporal processing

To explore whether the cortical representation of sound features is affected by the differences in subcortical coding fidelity, we acquired EEG data as listeners performed the envelope-ITD detection task (for the timing of the sequence of events constituting a trial, see Fig. 7A). Both the induced power (Fig. 7B) and the phase-locking (Fig. 7C) spectra show responses locked to the onset of the notched noise, which starts 400 ms before the target modulated tones. The late sustained responses in the induced power (after 1500 ms) are likely related to the motor preparation for the button press response. Although the transposed tones do not induce a change in power compared to baseline, there is a clear phase-locked onset response in the low frequencies (<20 Hz). In addition, both the variable and standard tones elicit ASRs phase locked to the 40 Hz stimulus modulation frequency (Fig. 7C; onsets at 0 and 0.8 s, respectively).

In general, the MEG/EEG onset response to the second (“probe”) of a pair of identical stimuli is smaller than to the first (the “adaptor”), a result typically attributed to adaptation (Jääskeläinen et al., 2004). When the adaptor differs from the probe in one or more attributes, the response to the probe is less adapted/suppressed (Jääskeläinen et al., 2004). We used this fact to test cortical sensitivity to changes in ITDs. We explored how the response to a standard probe tone with 0 ITD depended on the ITD of a preceding variable-ITD adaptor. To quantify the adaptation of cortical onset responses, we computed the aggregate of the phase-locked EEG response from 0 to 20 Hz. We did this separately for trials where the envelope ITD in the adaptor was large (400 and 800 μs; greater than the median threshold) versus when it was small (50, 100, and 200 μs; smaller than the median threshold; see Fig. 7D). We found that when the ITD of the adaptor was small (i.e., more similar to the probe), the response to the subsequent 0-ITD probe was smaller than when the ITD of the adaptor was large (i.e., more dissimilar to the probe; paired t test, $t = 2.8, p = 0.01$). This difference is a form of envelope-ITD-specific adaptation similar to what has been reported previously for stimuli with broadband ITDs (Salminen et al., 2010). For each individual, we computed the “adaptation ratio” by computing the ratio of the probe onset response when preceded by a large-ITD adaptor compared to a small-ITD adaptor. A weak but statistically significant negative correlation was found between the envelope-ITD threshold and the adaptation ratio across individuals ($r = −0.34, p = 0.04$). In other words,
listeners who have relatively poor sensitivity and large behavioral thresholds tend to have small onset responses to the standard tone preceded by a large-ITD tone, as if for them the large-ITD variable tones are perceptually similar to the 0-ITD standard tone. These results suggest that subcortical temporal coding fidelity varies across NH listeners, affecting both behavior and cortical physiology.

In contrast with the 100 Hz EFR, analysis of the 40 Hz cortical ASSR revealed that there were considerable across-subject differences in the background cortical activity around the 40 Hz region. This may be due to anatomical factors, task-related background cortical activity, as well as other factors. Thus, we did not pursue any further analysis of this response.

**Relationship between attention task performance and correlates of cochlear mechanics and suprathreshold temporal coding**

Listeners reported the content of one of two simultaneous streams of spoken, monotonized digit sequences based on direction, which was controlled by imposing different ITDs on the streams (Fig. 8A). Results revealed significant individual differences, with performance ranging from chance level to ceiling (Fig. 8B). Errors generally arose because listeners reported the masker instead of the target, not from lapses in intelligibility or memory; specifically, 91% of the incorrectly reported digits appeared in the masking stream at the same position in the sequence. This suggests that sensory encoding limitations made it difficult to spatially resolve the target and masker, leading to attentional selection errors. To evaluate the relative contributions of peripheral processing and neural temporal coding to individual differences in the spatial attention task, we entered the data into a generalized linear regression analysis. Performance (number of correct responses) was modeled as coming from a binomial distribution. The latent binomial parameter was modeled through a logit link function using a linear mixed model (Hadfield, 2010) with ITD, individual EFR slopes, and DPOAE thresholds as explanatory variables. Inference was performed using MCMC sampling from the posterior. The coefficient of the DPOAE term was not distinguishable from zero, whereas the EFR term was significant ($p = 0.0024$). The marginal posterior of the coefficient of the EFR term was unimodal, with a 95% highest-posterior-density interval (Box and Tiao, 2011) spanning $-5.9$ to $0.95\%$ with a mean of $-3.5\%$; that is, for every 1 dB/4 dB steepening of the EFR slope, the probability of a correct response dropped by an average of 3.5%. Not surprisingly, the effect of ITD was also significant ($p = 0.004$), suggesting, along with the error patterns, that sensory limitations were a key determinant of performance. Thus, for this particular spatial attention task, individual differences in temporal coding accounted for much of the variance in performance.

**Noise-exposure history**

To get a sense of the relationship between noise exposure and subcortical temporal coding, we asked our subjects to rate their own exposure on a closed scale. Exposure history was assessed by considering the frequency that listeners reported attending concerts, loud parties, and sports venues and using earphones. In addition, subjects indicated whether they recalled ever having temporary threshold shifts (“losing sensitivity to faint sounds for a while”) following noise exposure. Subjects were divided into two groups (“more exposed” and “less exposed”) based on their responses, with equal weight given to all categories of exposure. Compared to the less exposed group, the more exposed group had a significantly steeper EFR slope ($t_{(23)} = 1.92, p = 0.034$) along with higher AM ($t_{(23)} = 1.97, p = 0.03$) and envelope-ITD ($t_{(23)} = 2.35, p = 0.014$) thresholds compared to the less exposed group. These effects remain marginally significant when the false discovery rate for multiple comparisons is controlled at $5\%$ (Benjamini and Hochberg, 1995). Given the relatively crude characterization of noise-exposure history, correlations between all the temporal coding measures, and the size of our cohort, this result should be interpreted with caution.

**Discussion**

In the conventional view, hearing loss is associated with cochlear mechanical anomalies, typically caused by hair cell loss or damage to mechanosensory hair bundles (for a detailed treatment, see Moore, 2007). In contrast, previous animal studies showed that moderate noise exposure and early aging can produce synapse/neuronal loss without changes to cochlear mechanics and with no
hair cell loss—a completely different form of sensorineural hearing loss (Kujawa and Liberman, 2009; Makary et al., 2011; Sergeyenko et al., 2013).

Here, we investigated whether perceptual differences in the general population of NH listeners who have no known hearing deficits are consistent with this mechanism of synaptopathy/neuropathy (Bharadwaj et al., 2014). Our measures were designed to reveal differences in suprathreshold coding fidelity, using stimuli with relatively high sound levels, competing maskers, and subtle temporal features. These choices were important for two reasons: (1) temporal coding deficits that are not apparent in ordinary laboratory conditions may be exposed under such more challenging regimes (Fig. 1), and (2) low-SR fibers, which are especially vulnerable to neuropathy (Fig. 2) are likely to contribute more under such conditions. Specifically, compared to high-SR fibers, the low-SR ANFs are more resistant to masking (Costalupes, 1985), and exhibit more robust synchrony to both envelope and fine structure in high-intensity sounds (Johnson, 1980; Joris and Yin, 1992). To help isolate where differences arise, we evaluated peripheral transduction, early neural coding, cortical sensitivity to ITDs, and suprathreshold temporal perception. We found large variations in perceptual ability on both low-level psychophysical tasks and a high-level selective attention task; across subjects, these behavioral measures strongly correlated with physiological measures of suprathreshold temporal coding. Furthermore, we established that differences in cortical sensitivity of binaural information inherit some of the differences in the fidelity of early neural temporal coding. Crucially, independent measures of cochlear amplifier function, though correlated with each other, accounted for very little of the variability in attention task performance and were uncorrelated with psychophysical and physiological measures of suprathreshold temporal coding. Together,
Figure 7. Results from cortical EEG measures of envelope ITD sensitivity. 

**A**, Stimulus sequence. Subjects were asked to indicate if sounds jumped from left to center or right to center. 

**B**, Induced power relative to baseline. 

**C**, Intertrial phase locking reveals low-frequency responses to the onset of AM tones in addition to showing a sustained response at the modulation frequency (40 Hz). 

**D**, A closer look at the low-frequency power in onset responses to the two tones (averaged over frequencies below 20 Hz) showing envelope ITD–specific adaptation.
these results demonstrate that individual differences in NH listeners are in part driven by differences originating very early in the neural (and not mechanotransduction-related) portions of the auditory pathway.

To quantify physiological coding strength, we settled on the slope of the EFR as a function of modulation depth. Whereas absolute EFR strength (Fig. 6B) also correlated with behavioral performance, the EFR slope proved to be more a sensitive correlate of perceptual ability. This is likely both because the EFR slope (1) helps normalize differences in EFR magnitude due to individual differences in measurement noise, brain tissue geometry, and other phenomena leading to purely linear transformations of the responses, and (2) emphasizes coding fidelity of shallow stimulus modulations at high sound levels. The low level effect is consistent with the loss of high-SR/low-SR counts from 55/45 to 75/25 is consistent with some high-SR neuropathy being present (Bharadwaj et al., 2014). Consistent with the interpretation that differences are present as early as at the level of the auditory nerve, the EFR slope correlated with sensitivity to both monaural and interaural temporal information. Furthermore, the efficacy of the EFR slope in exposing individual differences is in accordance with a previous computational model of low-SR neuropathy (Bharadwaj et al., 2014). Of course, unlike in animal models, we cannot manipulate or measure cochlear neuropathy in humans directly; our interpretation is based on indirect evidence. Still, the observed pattern of results is consistent with that mechanism.

We found robust correlations across various measures of temporal coding, but no correlation between measures of cochlear mechanical function and perceptual abilities. Of course, if a larger cohort of NH participants were tested, residual differences in cochlear mechanics could well explain a significant portion of the individual variability. Regardless, our results raise the possibility that noise-induced neuropathy is a significant contributor to what has until now been viewed as “normal” variability in perceptual hearing ability. Indeed, an estimated 17% of the general population experience occupational noise at levels sufficient to produce temporary threshold shifts and have subsequent hearing difficulties that are attributable to this exposure (Tak and Calvert, 2008; Tak et al., 2009). Recreational noise exposure is more difficult to quantify; however, estimates suggest that exposures can often exceed 100 dB SPL (Clark, 1991). Previous evidence suggests that even moderate exposures over extended periods are neuropathic (Maison et al., 2013). Thus, in modern society, where noise exposure is common, it is plausible that interindividual variability in hearing ability may partly originate from hidden deficits caused by noise exposure or its interactions with aging and genetic susceptibility (Davis et al., 2003; Kujawa and Liberman, 2006). The correlations, albeit weak, between temporal coding measures and noise-exposure history support the notion that neuropathy may be present even in normal-hearing listeners. Recent independent findings relating noise-exposure to EFR measures are consistent with this interpretation (Plack et al., 2014). Stone and Moore (2014) found that higher levels of recreational noise exposure in young listeners were associated with poorer AM detection thresholds, but only at lower sound levels. The low level effect is consistent with the loss of high-SR fibers; indeed, an ~40% loss of overall synapse count (Lin et al., 2011a) accompanied by change in population statistics of high-SR/low-SR counts from 55/45 to ~75/25 is consistent with some high-SR neuropathy being present (Fig. 2). That Stone and Moore (2014) did not see noise-exposure effects at higher sound levels may possibly be due to associated increases in absolute thresholds that they found; outer hair cell loss can cause broadened tuning and loudness recruitment leading to improved AM detection at higher levels (Moore et al., 1996). Additional studies are needed to clarify these issues and assess the sensitivity of behavioral and EFR-based approaches in revealing noise damage when multiple mechanisms are involved simultaneously.

In real-world situations, selective attention performance depends on many nonsensory factors such as language ability,
memory, etc. (Conway et al., 2001; Surprenant and Watson, 2001; Kidd et al., 2007). However, in this study, we wanted to determine whether sensory coding fidelity of suprathreshold sound impacts the ability to communicate in everyday settings. Accordingly, we designed a spatial attention task where language and memory demands were low and sensory encoding limitations were the likely bottleneck on performance. Error patterns confirm that we achieved this goal. The fact that selective attention performance correlates with suprathreshold coding fidelity lends credence to the idea that neuropathy influences how well listeners can communicate in daily life. Although aging and noise exposure likely contribute to suprathreshold hearing deficits, inherent variability from genetic, epigenetic, and experiential factors no doubt also influence individual outcomes. Experience- and training-dependent long-term plasticity effects are known to affect processing at various stages of the auditory system and can improve behavioral performance in auditory tasks (Russo et al., 2005; Polley et al., 2006). Additional studies are needed to assess the contribution of such long-term plasticity to the differences in electrophysiological measures (such as the EFR) in the general NH population and to ascertain the degree to which such plasticity can ameliorate the effects of neuropathy. Individual differences in the olivocochlear efferent system and middle ear reflexes also must be examined (Backus and Guinan, 2007); however, the relationship between efferent effects and EFR measures is likely complicated given that (1) activation of the efferent system may help protect listeners from noise-induced cochlear neuropathy (Maison et al., 2013) and (2) low-SR nerve activity may be linked to the acoustic activation of the efferent system and the stapedius reflexes (Liberman, 1988; Kobler et al., 1992). Thus, once some degree of neurodegeneration occurs, it is possible that the efferent system may become less effective, leading to further damage. Additional studies are needed to understand the undoubtedly complex relationships between long-term plasticity, efferent function, middle-ear reflexes, and neuropathy.

We conclude that NH listeners from the general population vary significantly in how sensitive they are to suprathreshold temporal structure in sound. These differences appear to be unrelated to cochlear transduction, consistent with the insensitivity of the audiogram to such differences. Instead, results suggest that these differences come from the fidelity with which temporal features are encoded by very early levels of the neural pathway and are consistent with cochlear neuropathy degrading suprathreshold hearing ability in some listeners. These deficits may not be evident in quiet settings, but will affect how well listeners communicate in common social settings. In the field of hearing science, there is a growing belief that such hidden hearing loss is a unrecognized and poorly understood public-health concern. Hidden hearing loss has also been hypothesized to underlie tinnitus in listeners with normal audiograms (Schaette and McAlpine, 2011; Eggermont and Roberts, 2014; Plack et al., 2014). Thus there is tremendous interest in objective assessment tools for such deficits. Our results suggest that EFR-based measures are promising in this respect. Additional experiments are needed to develop targeted methods for the assessment, detailed characterization, and prevention of such hidden hearing loss.

References


Joris PX, Carney LH, Smith PH, Yin TC (1994) Enhancement of neural activity in the efferent cochlear nucleus: help protect listeners from noise-induced cochlear neuropathy that can ameliorate the effects of neuropathy. Individual differences in the olivocochlear efferent system and middle ear reflexes also must be examined (Backus and Guinan, 2007); however, the relationship between efferent effects and EFR measures is likely complicated given that (1) activation of the efferent system may help protect listeners from noise-induced cochlear neuropathy (Maison et al., 2013) and (2) low-SR nerve activity may be linked to the acoustic activation of the efferent system and the stapedius reflexes (Liberman, 1988; Kobler et al., 1992). Thus, once some degree of neurodegeneration occurs, it is possible that the efferent system may become less effective, leading to further damage. Additional studies are needed to understand the undoubtedly complex relationships between long-term plasticity, efferent function, middle-ear reflexes, and neuropathy.

We conclude that NH listeners from the general population vary significantly in how sensitive they are to suprathreshold temporal structure in sound. These differences appear to be unrelated to cochlear transduction, consistent with the insensitivity of the audiogram to such differences. Instead, results suggest that these differences come from the fidelity with which temporal features are encoded by very early levels of the neural pathway and are consistent with cochlear neuropathy degrading suprathreshold hearing ability in some listeners. These deficits may not be evident in quiet settings, but will affect how well listeners communicate in common social settings. In the field of hearing science, there is a growing belief that such hidden hearing loss is a under-recognized and poorly understood public-health concern. Hidden hearing loss has also been hypothesized to underlie tinnitus in listeners with normal audiograms (Schaette and McAlpine, 2011; Eggermont and Roberts, 2014; Plack et al., 2014). Thus there is tremendous interest in objective assessment tools for such deficits. Our results suggest that EFR-based measures are promising in this respect. Additional experiments are needed to develop targeted methods for the assessment, detailed characterization, and prevention of such hidden hearing loss.

References

