A Hearing-Model-Based Active-Learning Test for the Determination of Dead Regions

Josef Schlittenlacher[1], Richard E. Turner[2], and Brian C. J. Moore[1]

Abstract

This article describes a Bayesian active-learning procedure for estimating the edge frequency, f_e, of a dead region, that is, a region in the cochlea with no or very few functioning inner hair cells or neurons. The method is based on the psychophysical tuning curve (PTC) but estimates the shape of the PTC from the parameters of a hearing model, namely f_e, and degree of outer hair cell loss. It chooses the masker frequency and level for each trial to be highly informative about the model parameters in the context of previous data. The procedure was tested using 14 ears from eight subjects previously diagnosed with high-frequency dead regions. The estimates of f_e agreed well with estimates obtained using "Fast PTCs" or more extensive measurements from an earlier study. On average, 33 trials were needed for the estimate of f_e to fall and stay within 0.3 Cams of the final "true" value on the equivalent rectangular bandwidth-number scale. The time needed to obtain a reliable estimate was 5 to 8 min. This is comparable to the time required for Fast PTCs and short enough to be used when fitting a hearing aid. Compared with Fast PTCs, the new method has the advantage of using yes-no judgments rather than continuous Békésy tracking. This allows the slope of a subject's psychometric function and thus the reliability of his or her responses to be estimated, which in turn allows the test duration to be adjusted so as to achieve a given accuracy.

Introduction

Cochlear hearing loss is often associated with damage to the hair cells (Borg, Canlon, & Engström, 1995; Engström, 1983; Schuknecht & Woellner, 1953; see also Moore, 2007). Outer hair cell loss (OHCL) reduces the gain of the cochlear amplifier, while inner hair cell loss (IHCL) reduces the number of action potentials evoked in the auditory nerve. The total amount of cochlear hearing loss, HL_{total}, can be modeled as the sum of OHCL and IHCL, when all three are expressed in dB (Moore & Glasberg, 2004). Usually, OHCL is much larger than IHCL. In some cases, there may be a region in the cochlea where the inner hair cells or neurons are functioning so poorly that a pure tone producing peak vibration in that region is detected by off-place listening (also called off-frequency listening). In other words, the tone is detected at a place whose characteristic frequency (CF) is different from the frequency of the tone. Such a region is known as a *dead region* (DR; Moore, 2001, 2004). The most common type of DR is a basal DR, which starts at a place in the cochlea with

CF $= f_e$ and extends upwards from f_e. Diagnosing the presence of a basal DR, and estimating f_e is relevant to the fitting of hearing aids, including cases where a hearing aid is combined with a cochlear implant (Moore, 2001; Moore & Malicka, 2013; Zhang, Dorman, Gifford, & Moore, 2014). It may also be relevant to deciding whether a hearing-impaired (HI) person is a candidate for a cochlear implant (Moore, Glasberg, & Schlueter, 2010). This article describes a novel, behavioral method of estimating f_e that is both time-efficient and accurate.

As described earlier, a pure tone whose frequency falls in a DR may be heard because of off-frequency listening: The excitation evoked in the cochlea spreads to CFs

[1]Department of Experimental Psychology, University of Cambridge, UK
[2]Department of Engineering, University of Cambridge, UK

Corresponding author:
Josef Schlittenlacher, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 2EB, UK.
Email: js2251@cam.ac.uk

away from the signal frequency, and if the tone is sufficiently intense, this can lead to audible excitation in an adjacent non-DR. This makes exact determination of f_e difficult. Even if there is a sharp boundary between functioning and nonfunctioning inner hair cells, the threshold for detecting a tone may increase only gradually as its frequency traverses the boundary.

A well-established test for DRs is the TEN(HL) test (Moore, Glasberg, & Stone, 2004). The task is to detect a pure-tone signal in a broadband masker, threshold-equalizing noise (TEN). The TEN has approximately a constant level per ERB_N, where ERB_N is the equivalent rectangular bandwidth of the auditory filter for listeners with normal hearing (Glasberg & Moore, 1990). The masked threshold, expressed as the signal level relative to the level of the TEN in a 1-ERB_N-wide band at 1 kHz (denoted the signal-to-noise ratio [SNR]), is about 0 dB for listeners with normal hearing, that is, the signal needs to have the same level as the TEN produces at the output of the auditory filter centered at the signal frequency. For HI listeners without DRs, the threshold for detecting the signal in the TEN is typically about 3 dB higher. However, if the frequency of the signal falls well inside a DR, it will only be detected if it evokes sufficient excitation in the non-DR below f_e, and this usually requires an SNR of 10 dB or more (Moore et al., 2004; Moore, Huss, Vickers, Glasberg, & Alcántara, 2000). Hence, a DR at the signal frequency is diagnosed when the SNR at threshold is 10 dB or more (with the additional condition that the signal level at the masked threshold should be at least 10 dB above the absolute threshold). The TEN test allows rapid detection of a DR, but it does not provide a precise estimate of f_e. For example, for a basal DR, if the frequency of the tone is only a little above f_e, the SNR at threshold may be only a little above the "normal" value of 0 dB (Moore, 2001, 2004), and the DR will be "missed". For a basal DR, the "true" value of f_e usually lies somewhat below the lowest signal frequency at which the TEN-test criteria are met.

Another test for DRs is based on psychophysical tuning curves (PTCs; Chistovich, 1957). Here, the task is to detect a pure-tone signal in the presence of a narrowband noise. The frequency (f_{sig}) and level (L_{sig}) of the signal are fixed, and the center frequency and level of noise (f_{mask}, L_{mask}) are varied. A PTC is the value of L_{mask} required to mask the signal plotted as a function of f_{mask}. For normally hearing subjects and subjects with hearing loss but without DR, the PTC has a minimum near f_{sig}, that is, where the frequencies of the signal and noise coincide. When the signal frequency falls within a DR, the noise only needs to mask the part of the excitation pattern that falls outside the DR. Thus, the minimum in the PTC occurs when $f_{mask} \approx f_e$. PTCs measured using several discrete values of f_{mask} have been used to determine f_e in experimental studies (Kluk

& Moore, 2005; Moore & Alcántara, 2001; Moore et al., 2000). However, this method is too time-consuming for use in clinical practice.

A method of obtaining PTCs more rapidly is called "Fast PTCs" (Sęk, Alcántara, Moore, Kluk, & Wicher, 2005). The method uses Békésy tracking: f_{sig} and L_{sig} are kept constant, and the signal is continuously pulsed on and off, while f_{mask} is continuously increased (or decreased) over time. The subject presses a button to indicate whether he or she hears the signal, and L_{mask} is increased while the button is pressed and decreased while it is released. This yields a PTC within 3 to 6 min, depending on the selected range of f_{mask} and the rate of change of f_{mask}. However, the method has the disadvantage that the subject may "lose track" of the signal; essentially, the subject forgets what to listen for. For some subjects, many Fast PTCs using several values of f_{sig} and L_{sig} are needed to obtain a robust estimate of f_e (Kluk & Moore, 2005, 2006).

The present article describes a PTC-based method that is fast but also accurate. The masked threshold of a pure tone in noise, and thus a PTC, is predicted using a hearing model (Moore & Glasberg, 2004) based on the simple assumption that the signal can be heard if its excitation exceeds the excitation evoked by the noise at any CF and cannot be heard if its excitation is below the excitation evoked by the noise or below the audiometric threshold at all center frequencies. The calculated excitation patterns of the tone and noise depend on their frequency and level and the subject's hearing loss. The model parameters used to characterize the hearing loss are HL_{total}, OHCL, and f_e. HL_{total} is known from the audiogram. In practice, when f_{sig} falls above f_e, the maximum excitation evoked by the signal always occurs at f_e and the highest ratio of signal excitation to masker excitation occurs at f_e. Therefore, the value of OHCL only needs to be estimated at f_e; this value is denoted OHCL(f_e). Responses obtained using various combinations of f_{mask} and L_{mask} limit the possible combinations of f_e and OHCL(f_e).

In the proposed method, a probabilistic approach is used to express the degree of "belief" about the values of f_e and OHCL(f_e). The belief is represented by a probability distribution over f_e and OHCL(f_e), $p(f_e, \text{OHCL}(f_e))$, that indicates how likely each setting of f_e and OHCL(f_e) is, given the data. As the test proceeds, $p(f_e, \text{OHCL}(f_e))$ collapses to a single point, that of the true parameters. To make the test fast, f_{mask} and L_{mask} are chosen to be highly informative about f_e and OHCL(f_e). A similar approach was used by Kontsevich and Tyler (1999) to estimate a psychometric function, its mean and slope being the model parameters. For each trial, the stimulus level was chosen to minimize the expected entropy (Shannon, 1948) of the model parameters given the previous data (see the Method section for details). The psychometric function has only a single

stimulus parameter, the level. Other work in cognitive science has made this strategy tractable when two or more parameters need to be estimated (DiMattina, 2015; Houlsby, Huszár, Ghahramani, & Lengyel, 2011; Houlsby et al., 2013; Kujala & Lukka, 2006), by using equivalent but computationally less expensive formulae. Active-learning paradigms that maximize the information gained on each trial have been used in the auditory domain to estimate the bandwidth of the auditory filter (Shen & Richards, 2013) and to estimate the audiogram (Song, Sukesan, & Barbour, 2018; Song et al., 2015). In this article, we show that the proposed Bayesian active-learning procedure determines f_e accurately within minutes. It is thus short enough to be applied in clinical practice.

Method

Overview

First, two audiograms were obtained for each test ear of each subject. One was obtained using an audiometer and the manual procedure recommended by the British Society of Audiology (2011). The other was obtained using an active-learning procedure similar to that described by Song et al. (2015). On average, the threshold obtained using the active-learning method was 3.2 dB below that obtained using the audiometer (standard deviation of the difference: 6.7 dB). The variability is comparable to that reported by Song et al. (2015). The audiograms obtained using the active-learning procedure were used for further analysis since they were obtained on the same apparatus and yielded a continuous estimate of the threshold as a function of frequency.

It has been shown that the likelihood of a DR being present at a given frequency increases markedly when HL_{total} at that frequency exceeds 65 dB HL (Vinay & Moore, 2007). Hence, in a second step, the frequencies at which the audiometric threshold reached 65, 70, and 75 dB HL (f_{65}, f_{70}, and f_{75}) were extracted, and a quick version of the TEN(HL) test was used to assess whether f_{65} fell in a DR (see later for details). If it did not, the quick TEN(HL) test was repeated to assess whether f_{70} and f_{75} fell in a DR. Finally, each subject was tested using a Bayesian active-learning procedure (called "Smart DRT") to estimate f_e, and a Fast PTC was also obtained. Since the frequency at the tip of the Fast PTC is usually taken as an estimate of f_e, one can think of the goal of the Smart DRT as being to estimate the frequency at the tip of the PTC but without determining the entire PTC. To assess the repeatability of the Smart DRT, it was administered three times for each test ear. Half of the ears were tested using the Smart DRT before the Fast PTC, and the other half were tested in the reverse order.

For all tests, the subject was asked to indicate whether the signal was present on each trial or not, that is, a yes-no task was used. This was chosen in preference to a two-interval forced-choice task since random guessing in a two-interval forced-choice means that on average, a correct response will be given on half of trials where the signal was not heard at all, whereas a "yes" response in a yes-no task probably indicates that the subject did obtain some sensory evidence indicating the presence of the signal and is therefore more informative (Green & Sweets, 1974). In most cases, subjects adopted a cautious response criterion, usually responding "no" on trials where the signal was absent.
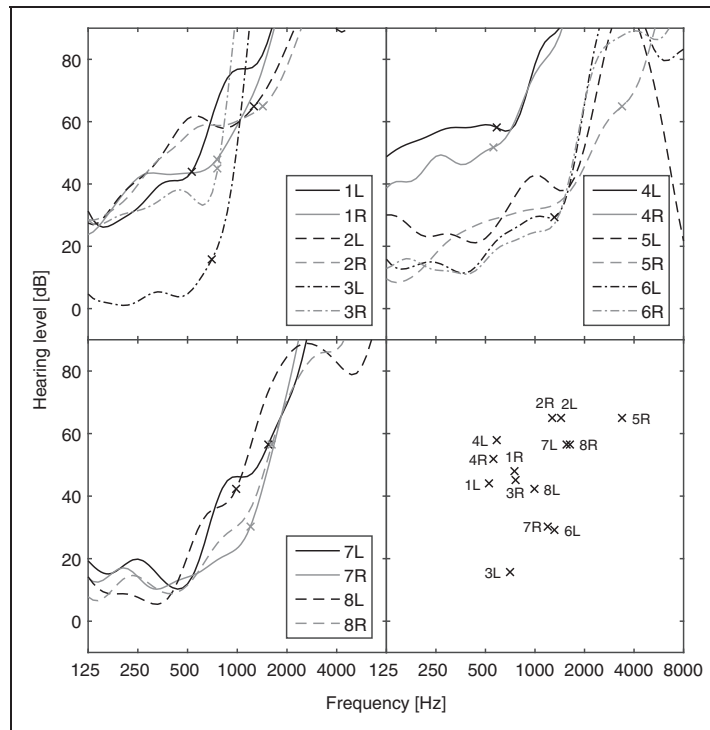
Subjects

Fourteen ears from seven male subjects and one female subject aged 45 to 82 years (mean 72 years) were tested. All test ears had previously been diagnosed as having extensive basal DRs using the TEN(HL) test and Fast PTCs, as described by Salorio-Corbetto, Baer, and Moore (2017). Six of the subjects had DRs in both ears, and two had a DR in one ear only. The contralateral ears of these two subjects had impaired hearing but no DR. All subjects had normal hearing to moderate hearing loss at low frequencies and hearing loss of more than 65 dB at high frequencies. The slope of the hearing loss varied across subjects. At a hearing loss of 65 dB HL, the derivative of HL_{total} ranged from 14 dB/octave to 145 dB/octave. The audiograms for all test ears, including the two without a DR, are shown in Figure 1, with lines denoting HL_{total} and crosses the values of f_e that were determined by the Smart DRT. In addition, the quick TEN(HL) test was conducted using two normal-hearing and two HI subjects without a DR, in order to confirm that the quick TEN(HL) test did not falsely identify DRs when none were present.

Apparatus

For all tests except standard audiometry, stimuli were generated digitally at a sample rate of 48 kHz using a personal computer. The signal was converted to analog form by an M-Audio Delta 44 audio interface (Cumberland, RI) and presented via a Sennheiser HDA200 headset (Wedemark, Germany). The frequency response of the headset at the eardrum was estimated using KEMAR (Burkhard & Sachs, 1975) and used to determine sound levels at the eardrum. The traditional audiogram was measured using a Grason-Stadler GSI-61 audiometer and Telephonics TDH-50 headphones.

Stimuli

All signals were presented monaurally. For the quick TEN(HL) test, a TEN(HL) noise was generated in the same way as described by Moore et al. (2004), but it was

Figure 1. Audiograms obtained using the active-learning method for the 14 test ears and the two contralateral ears without DR of the same subjects (5L and 6R). Lines show HL$_{\text{total}}$, and crosses show the values of f_{e} that were determined by the Smart DRT. The two upper panels and the lower left panel show results for subsets of ears, and the bottom right panel shows the values of f_{e} plotted against the audiometric threshold at f_{e} for all ears with DRs.

designed for use with Sennheiser HDA200 headphones. It was band limited between 354 and 6500 Hz, and its level is specified in dB HL/ERB$_{\text{N}}$ at 1 kHz. The noise had a low crest factor, to allow presentation at high levels without peak clipping. However, after passing through the auditory filters, the noise would have had a Gaussian distribution of instantaneous amplitudes.

The task in the Smart DRT was similar to that for a traditional PTC, that is, to detect a pure tone in the presence of bandpass-filtered noise. The bandwidth of the masking noise was 1 ERB$_{\text{N}}$ or 200 Hz, whichever was greater. The bandwidth was selected to minimize the influence of beats on the masked threshold (Kluk & Moore, 2004, 2005). The level (L_{mask}) and center frequency (f_{mask}) of the noise were varied. The noise had a flat spectrum within the passband and was generated using an inverse discrete Fourier transform, the spectrum being defined with a resolution of 1 Hz. A broadband TEN with a level of 35 dB HL/ERB$_{\text{N}}$ was added to mask potential distortion products (Alcántara & Moore, 2002).

The temporal envelopes of the stimuli were the same for the quick TEN(HL) test and the Smart DRT. The signal consisted of three pulses of a tone with Hann-windowed rise and fall times of 20 ms and a total duration of 150 ms each. The interval between pulses was 100 ms.

The noise was switched on 100 ms before the first pulse started and switched off 100 ms after the last pulse finished, again with Hann-windowed rise and fall times of 20 ms, leading to a total duration of 850 ms. The stimuli for determining Fast PTCs were generated using the software described by Sęk and Moore (2011).

Procedure for the Quick TEN(HL) Test

The quick TEN(HL) test was conducted to determine whether the test ear had a DR at f_{65}, f_{70}, or f_{75}. The value of f_{sig} was initially set equal to f_{65}. L_{sig} was set to 75 dB HL, which is 10 dB above the absolute threshold determined using the active-learning procedure, that is, 10 dB SL. The TEN level was 65 dB HL/ERB$_{\text{N}}$. This slightly lower TEN level than that proposed by Moore et al. (2004) was used here to avoid uncomfortably loud levels, especially for subjects with good low-frequency hearing. Initially, the signal was presented without any noise, and the subject was asked to confirm that it was audible. Thereafter, 10 trials containing the signal in TEN and 10 trials with the TEN only were presented in random order. The task was to indicate whether the signal was heard or not. Denote the proportion of trials where the signal was present and the subject responded "yes" (hits) as X and the proportion of trials where the

signal was absent and the subject responded "no" (correct rejections) as Y. The proportion of false alarms (responses of "yes" when the signal was absent) is $1 - Y$. If the signal is not detectable at all, $X = 1 - Y$. We chose as a threshold criterion $X \geqslant 1.5(1 - Y)$, that is, $X + Y \geqslant 1.5$. Given that there were 10 trials of each type, this corresponds to a total number correct $\geqslant 15$. In practice, the scores were mostly close to 10 correct or (in the case of the subjects without DRs at f_{65}) close to 20 correct. Thus, the outcome was not affected by the exact threshold criterion used.

If the subject scored 14 or fewer correct, it was assumed that there was a DR at f_{65} (see Moore et al., 2004). If the subject scored 15 or more correct, the quick TEN(HL) test was repeated with f_{sig} set to f_{70}, L_{sig} set to 80 dB HL, and the TEN(HL) level set to 70 dB HL/ERB$_N$. If the subject scored fewer than 15, it was assumed that there was a DR at f_{70}. If the subject scored 15 or more, the quick TEN(HL) test was repeated with f_{sig} set to f_{75}, L_{sig} set to 85 dB HL, and the TEN(HL) level set to 75 dB HL/ERB$_N$. The signal was never detected in this case, indicating the presence of a DR at f_{75}. The value of f_{sig} used for the Smart DRT and Fast PTCs was taken as the lowest frequency (out of f_{65}, f_{70}, and f_{75}) at which the subject scored less than 15 correct. The signal level was always 10 dB SL.

Procedure for the Smart DRT

A single trial in the Smart DRT included three intervals. The task was to indicate whether a pure-tone signal was present in the third interval or not. For four practice trials and 100 subsequent trials, the intervals contained the signal alone, the noise alone, and the signal plus the noise, in that order. For 20 trials that were randomly mixed with the other 100 trials, the signal was omitted from the third interval, to assess false positives.

The procedure started with four easy practice trials to introduce the subject to maskers with different frequencies and levels. The experimenter inspected the results and could intervene if the subject performed unexpectedly badly. These four trials were not considered for any further analysis. The next 16 trials were chosen by simple rules: Initially, f_{mask} was set 2 Cams below f_{sig}, where Cams are units of the ERB$_N$-number scale (Moore, 2012). L_{mask} was increased across successive trials, with values of -20, -9, -6, -3, 0, 3, 6, and 9 dB relative to L_{sig}. This was then repeated with f_{mask} set 4 Cams below f_{sig}. This initial grid was intended to familiarize the subject with trials of varying detectability. Subsequent values of L_{mask} and f_{mask} were chosen by a model-based active-learning method. The responses for the initial 16 trials were taken into account by the active-learning method. Doing so limited the impact of a wrong response in the early active-learning trials.

There were two free parameters in the model. The value of f_e was one of them. It was assumed that when f_{sig} fell within a DR, detection of the signal would depend on the relative response to the masker and signal at the output of the auditory filter tuned to f_e, that is, on the relative excitation levels evoked by the signal and masker at f_e. The output of this filter was calculated using the method described by Moore and Glasberg (2004). The sharpness of the filter was assumed to be determined by the total hearing loss at f_e, denoted HL$_{total}(f_e)$, the amount of hearing loss at f_e that was attributed to reduced outer hair cell function, OHCL(f_e), and the spectrum of the input. The value of HL$_{total}(f_e)$ was based on the audiogram obtained using the active-learning procedure. The value of OHCL(f_e) was a second free parameter.

For given values of f_e and OHCL(f_e), the excitation levels evoked by the noise (E_{noise}) and by the signal (E_{sig}) were calculated. E_{noise} was substituted by the maximum of E_{noise} and HL$_{total}(f_e)$, because the threshold in quiet could limit the audibility of the signal. The difference between E_{sig} and E_{noise} was taken as an indicator of the audibility of the signal. To provide a very simple model for the audibility of the signal in noise, it was assumed that when the difference was 0 dB ($E_{sig} = E_{noise}$), the probability of a correct yes response (a hit, denoted $y = 1$) was 50%. The probability of a hit was calculated from a Gaussian cumulative density function (CDF) with a standard deviation of 3 dB

$$p(y = 1|x, \theta) = CDF\big(E_{sig}(\theta) - E_{noise}(\theta, x), \; \mu = 0, \; \sigma = 3\big) \quad (1)$$

where x denotes the presentation parameters f_{mask} and L_{mask}, and θ denotes the model parameters f_e and OHCL(f_e). A standard deviation of 3 dB was thought to be a reasonable estimate of the slope of the psychometric function for detection of a tone in noise (Green, Birdsall, & Tanner, 1957). In general, if the standard deviation is assumed to be smaller than it actually is, the algorithm becomes "overconfident", that is, assigns probabilities that are too high around the currently most probable model parameters. Without additional measures, this may result in the algorithm getting "stuck" at an incorrect estimate. If, on the other hand, the standard deviation is assumed to be larger than it actually is, the algorithm becomes more "cautious" and may waste time exploring extreme parameter values. The CDF was scaled to have values between 0.01 and 0.99 to incorporate a lapse rate for accidentally pressing the "wrong" button. Doing this was thought to allow the method to recover from "wrong" responses more quickly.

The probability of obtaining the observed sequence of yes and no responses after N trials, given a set of

model parameters and the presented stimuli, is as follows:

$$p(D|x,\theta) = \prod_{i=1}^{N} p(y_i|x_i,\theta) \qquad (2)$$

where D denotes the collected response data, and i denotes the ith trial. However, we are not interested in the probability of the collected data given some assumed model parameters, but in the probability of the model parameters, in particular f_e, given the collected data. The probability of a set of model parameters given the collected data can be obtained by Bayes' rule and is as follows:

$$p(\theta|D,x) \propto p(D|\theta,x)\,p(\theta) \qquad (3)$$

where $p(\theta)$ is the prior distribution of the model parameters, which does not depend on x. A uniform distribution was assumed for $p(\theta)$ with the constraint that $OHCL(f_e)$ could not be greater than $HL_{total}(f_e)$. Also, the sets $(f_e, OHCL(f_e))$ were restricted to avoid the entire audible region of the excitation pattern evoked by the signal falling above f_e. If this were the case, the signal would not even be audible in quiet according to the model, but it had been established that the signal in quiet with a level of 10 dB SL was clearly audible.

The parameters x^* of the next trial were chosen with a probability proportional to the mutual information I between the response on the next trial y^* and the model parameters θ,

$$I(y^*; \theta|x^*) = H(y^*|x^*, D) - E_{\theta \sim p(\theta|D)}(H[y^*|x^*, \theta]) \qquad (4)$$

where H is the Shannon entropy (Shannon, 1948), a measure of the uncertainty in a variable. The first term is the entropy of the expected response, and the second term is the expected conditional entropy of the response given the belief about the model parameters (for more details, see Houlsby et al., 2011). Mutual information is a mathematically principled measure of the mutual dependence between variables that underpins the field of information theory, and it is used here to pick stimulus parameters. Stimuli are informative about the parameters f_e and $OHCL(f_e)$ if the outcome is uncertain, that is, when the probability of responding "yes" is close to 50% rather than close to 0 or 100% (first term of Equation 4). In our case, this term is large when the masker level is close to the threshold value. The second term is low in regions of x^* where the responses have been inconsistent or where no data were collected. The net result is to explore stimuli that are close to the current threshold estimate, but distant from areas where "this" threshold estimate is confident.

Mutual information is an optimal criterion, but it can only be tractably optimized in a "greedy" manner, that is, the algorithm looks only one trial ahead and can only pick the next stimulus without considering potentially informative sets of several stimuli. Greedily selecting the stimulus that is most informative about θ can, in some circumstances, lead to failure to explore relevant regions of parameter space. Here, the failure would have been a result of the algorithm being unable to take into account the information gained on subsequent trials. This problem was avoided by picking the next stimulus with a probability proportional to the mutual information. This strategy leads to more exploration of the parameter space but still picks stimuli that are highly informative about θ.

In addition, the algorithm was "encouraged" to choose f_{mask} close to the current estimate of the most likely value of f_e, denoted $\hat{f}_e$, by applying an additional weight to f_{mask}, based on a normal distribution around $\hat{f}_e$ with a standard deviation of 1.5 Cams. This was done for several reasons. First, it avoided values of f_{mask} for which the model predictions were likely to be less accurate. Second, it utilized the fact that the PTC shape is most distinctive around f_e. Third, it favored low and thus more comfortable masker levels. Thus, the parameters for the next trial were chosen randomly with probability.

$$p(x^* = f_{mask}, L_{mask}) \propto I(y^*; \theta|x^*)N\left(x^*; \hat{f}_e, \sigma^2\right) \qquad (5)$$

where $\sigma = 1.5$ Cams and $N(\ldots)$ denotes the normal distribution. In addition, values of f_{mask} within 2 Cams of the value of f_{mask} on the previous trial were excluded, to avoid "duplication" of information.

The belief about f_e and $OHCL(f_e)$ was updated after each trial (Equation 3), and the $\hat{f}_e$ value that was most likely after the last trial was chosen as the final estimate of f_e. The computations were made tractable, so that they could be performed in the intertrial interval of about 2 s, by using numerical integration and precomputing excitation patterns before the experiment started.

Fast PTCs

Fast PTCs were obtained for comparison with the Smart DRT results, using the software of Sęk and Moore (2011). The values of f_{sig} and L_{sig} were those determined in the quick TEN(HL) test. The masker center frequencies covered the range from one octave below f_{sig} to slightly above f_{sig}. A low-pass noise with a cutoff frequency of 200 Hz and level of 40 dB SPL was added to mask possible difference tones (Alcántara & Moore, 2002). Two PTCs were obtained, one using an upward frequency sweep and one using a downward sweep with the masker level changing at a rate of 2 dB/s. The duration of each frequency sweep was 4 min. If a subject reported difficulties in the first run, for example

forgetting what to listen for, or if the results were erratic, for example because the PTC did not have a convex shape or clearly showed a lapse of attention, the run was repeated. In this case, only the results from the repeated run were analyzed.

Results

Quick TEN(HL) Test

The two normal-hearing subjects and two HI subjects without a DR were tested on the quick TEN(HL) test using $f_{sig} = 1\,kHz$, $L_{sig} = 75\,dB$ HL, and TEN(HL) level $= 65\,dB$ HL/ERB_N. All scored at least 19 out of 20 trials correct. For 13 of the test ears with DRs, the signal with frequency f_{65} was never heard when presented at 75 dB HL in the presence of the TEN(HL) at 65 dB HL/ERB_N, suggesting the presence of a DR at f_{65}. For the remaining test ear (3R), the subject achieved 20 and 19 trials correct for the frequencies f_{65} and f_{70}, but could not hear the signal with frequency f_{75} when presented at 85 dB HL in the presence of the TEN(HL) at 75 dB HL/ ERB_N, suggesting the presence of a DR at f_{75}.

Smart DRT

To illustrate the operation of the Smart DRT, Figure 2 shows the belief about f_e after 0, 20, 50, and 100 trials for the second run for ear 3 R. The probability of f_e was obtained by marginalizing the joint distribution over OHCL(f_e), that is, integrating the joint probabilities

over all OHCL(f_e) for each f_e. This averages away the uncertainty in OHCL(f_e). The prior probability (dotted line), which was estimated before considering any results for the initial grid, but just taking the audiogram and model constraints into account, was approximately uniform between about 11 Cams (520 Hz) and 14.8 Cams (900 Hz, corresponding to f_{sig}), and decreased to 0 at 7.2 Cams (270 Hz) and above 14.8 Cams (since it was already known that f_e was below f_{sig}). The distribution narrowed with increasing number of trials, N, with maxima at 13.3 Cams (730 Hz) after 20 trials and at 13.5 Cams (750 Hz) after 50 and 100 trials. On average, 33 trials were needed for the estimate of f_e to fall and stay within 0.3 Cams of the final estimate, indicating that less than 20 active-learning trials after the initial grid were sufficient to produce a reliable result (see also Figure 6).

The final estimates of f_e (those obtained at the end of each run) varied across runs of the Smart DRT and across ears from 0.5 kHz (ear 1 L) to 3.3 kHz (ear 5R). To assess the accuracy of the estimates, initially the values of f_e derived from the Fast PTCs for each ear were used as a "gold standard" or "reference". For the Fast PTCs, the value of f_e is usually taken as the frequency at the tip of the PTC, f_{tip} (Moore & Alcántara, 2001; Moore et al., 2000). However, the exact value of f_{tip} is sometimes difficult to estimate because of the jagged nature and somewhat irregular shapes of the Fast PTCs. Sęk and Moore (2011) proposed several methods for estimating f_{tip}, which can lead to somewhat different values. Here, the reference values of f_e were calculated by interpolating the 4-point moving averages
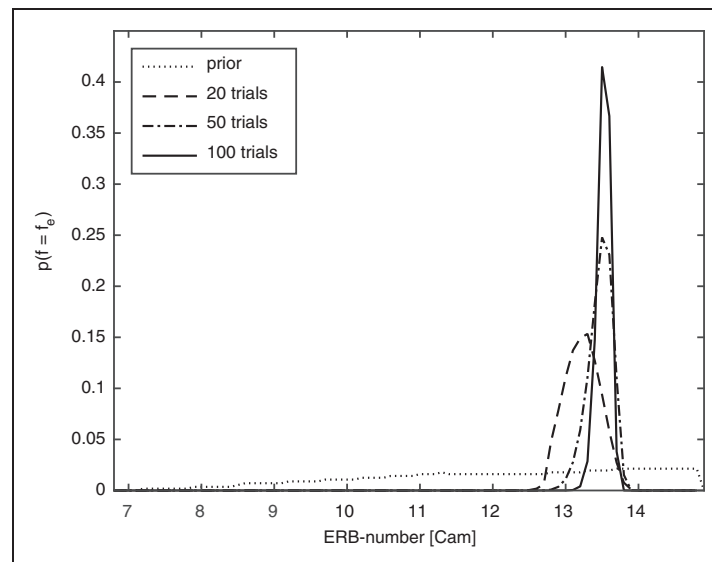

Figure 2. Probability density functions of the estimate of f_e, using a bin width of 0.1 Cam, for the second run of ear 3R. Dotted, dashed, dash-dotted, and solid lines show the probability density functions after 0, 20, 50, and 100 trials, respectively. Catch trials are not included in the number of trials.

of the Fast PTCs with splines, averaging the fitted levels at each masker center frequency for the upward sweep and the downward sweep, and taking the minimum of the curve obtained in this way. The Fast PTCs for 7L, 8L, and 8R did not show clear minima, and for these ears, the mean of the estimates of f_e for the three Smart DRT runs was taken as the reference. The Fast PTCs for 7L were very flat, with the downward sweep even being slightly concave. The Fast PTCs for 8L and 8R showed one or two reversals at the start for both the downward and upward sweeps but reached the maximum masker level immediately afterwards. Example Fast PTCs for 4L and 5R are shown in Figure 3. Both PTCs were rather flat, which is typical for ears with high values of OHCL. The irregular nature of the PTCs limits the accuracy with which f_{tip} can be determined.

For each test ear, the values of f_e obtained using each run of the Smart DRT were divided by the reference value of f_e for that ear. The resultant relative values of f_e are shown in Figure 4. Results for the three runs of the Smart DRT are indicated by circles, squares, and triangles. Ideally, these values should lie close to 1. In practice, they fell in the range 0.83 to 1.49 with a geometric mean of 1.06. However, the three Smart DRT estimates for each ear generally fell close to one another (except perhaps for 7L and 8L), indicating good internal consistency of the Smart DRT procedure. The range of the three Smart DRT estimates, expressed as the maximum estimate divided by the minimum estimate, had a geometric mean across ears of 1.08, and varied from 1.00 for 2L and 2R to 1.35 for 8L.

Two further estimates of f_{tip} were derived from the Fast PTC data (neither of which could be used with 7L, 8L, and 8R). For one, a 4-point moving average was obtained separately for the upward sweep and the downward sweep, the frequency at the minimum of each curve was found, and the two frequencies were averaged. The resulting estimates, expressed relative to the reference values of f_e, are shown as crosses in Figure 4. For the other additional estimate, a quadratic function was fitted separately to the data for the upward sweep and the downward sweep, the frequency at the minimum of each function was found, and the two frequencies were averaged. The resulting relative estimates are shown as × symbols in Figure 4. If the estimates of f_{tip} were consistent across methods, the crosses and × symbols would all fall close to 1. In fact, they fell between about 0.8 and 1.4. The variability across the three estimates of f_{tip} derived from the Fast PTC data was generally larger than the variability across the three estimates of f_e obtained from the three runs of the Smart DRT, suggesting that at least part of the deviation of the relative values from 1 for the latter arose from errors in the reference values. The range of the three Fast PTC estimates, expressed as the maximum estimate divided by the minimum estimate, had a geometric mean across ears of 1.19, and varied from 1.05 for 6L to 1.39 for 2L.

Some of the subjects in the present study had been extensively tested for DRs by Salorio-Corbetto et al. (2017). They performed a full TEN(HL) test and obtained Fast PTCs using several signal levels and frequencies. The estimates of f_e obtained by them (plotted
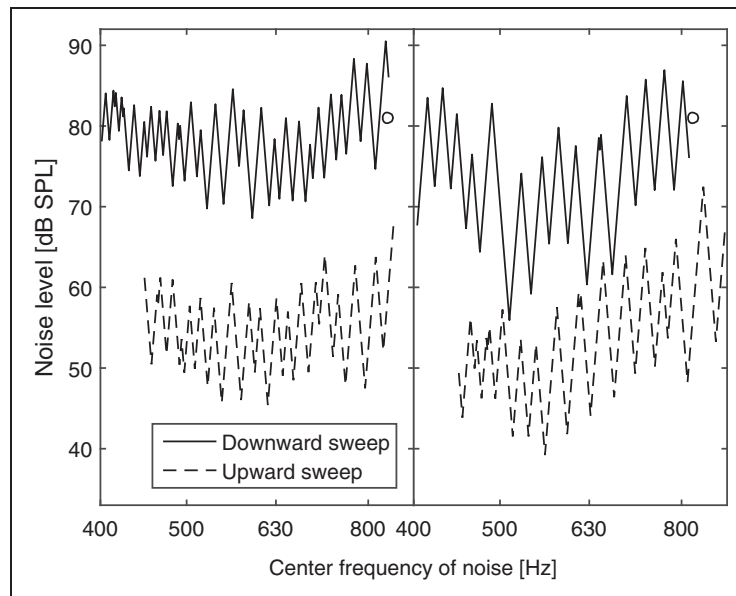
Figure 3. Fast PTCs for ears 4L (left) and 5R (right). The solid lines show PTCs for downward sweeps, and the dashed lines show PTCs for upward sweeps; the latter were shifted down by 25 dB for clarity. Circles show the frequency and level of the signal.
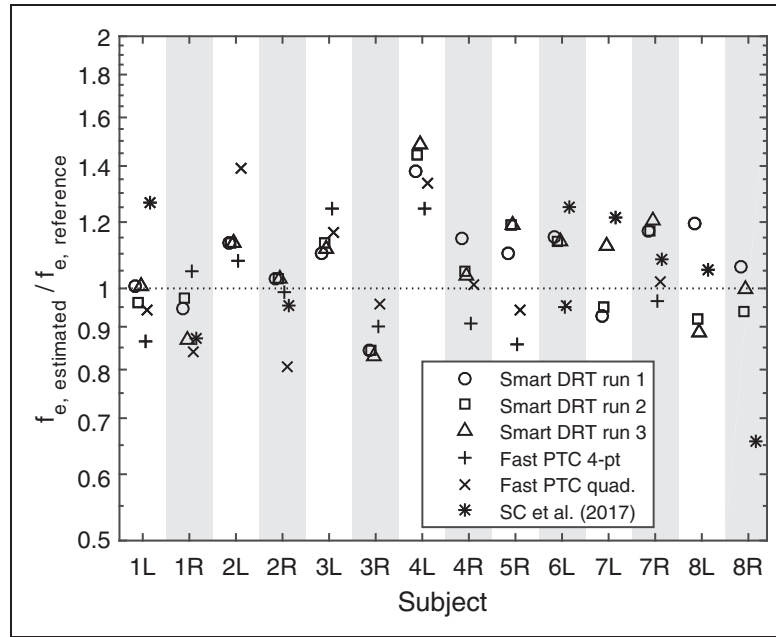
Figure 4. Estimated values of f_e expressed relative to a reference value, which was derived in most cases from the Fast PTCs (see text for details). Circles, squares, and triangles represent results for the three Smart DRT runs. Crosses and diagonal crosses represent two alternative measures of the tip frequency of the Fast PTCs, and asterisks show estimates obtained by Salorio-Corbetto et al. (2017) for some of the same ears as tested here.

relative to the reference values estimated here) are shown as asterisks in Figure 4. As noted earlier, in the present study, no estimates of f_{tip} could be obtained for ears 7L, 8L, and 8R. Salorio-Corbetto et al. (2017) reported that more runs than usual were needed for these ears to give stable results, but they were able to obtain Fast PTCs with clear minima. For ears 1R, 2R, 6L, 7R, and 8L, the estimates of f_{tip} obtained by Salorio-Corbetto et al. fell close to those obtained using the Smart DRT. However, for ears 1L, 7L, and 8R, there were marked deviations. For 1L and 7L, the estimates of Salorio-Corbetto et al. fell above those obtained here. This may have happened because they tested these ears 4 to 5 years ago, and the values of f_e may have shifted downwards over time.

Sixty catch trials (3 runs × 20 catch trials) were conducted for each ear. For 12 ears, the false positive rate was 3.3% or smaller, being 1.0% on average. Ears 5R and 1R gave false positive rates of 10% and 17%, respectively. Both subjects reported having tinnitus in their right ears.

It is useful to be able to check the consistency of the performance of a subject while the data are being collected. This can allow the test duration to be adjusted individually in order to achieve a predetermined level of accuracy. A possible method for doing this is to estimate the steepness of the psychometric function relating the proportion of "yes" responses to the sound level of the masker. This was done by expressing all masker levels relative to the PTC that was calculated from the most

likely combination of f_e and $OHCL(f_e)$. Having removed the effect of frequency in this way, the hyperparameters of a Gaussian Process with a linear kernel in level and Gaussian CDF likelihood function were optimized (see Rasmussen & Williams, 2006). Positive responses for levels higher than 3 standard deviations above the normalized masker level and negative responses for levels lower than 3 standard deviations below the normalized masker level were considered as outliers and discarded before the optimization was repeated iteratively. Figure 5 shows the resulting estimates of the standard deviation of the psychometric function for each subject, with the bars depicting the median for the three runs and the error bars depicting the range. White bars show estimates after 30 trials, and gray bars show estimates after all 100 trials. In some cases, the estimates after 30 trials were much lower than after all 100 trials or even close to zero. This may happen when the estimated PTC perfectly separates the audible and inaudible trials, as was the case for ears 3L and 3R. For the majority of ears, the estimated standard deviation was between 2 and 4 dB and thus close to the 3 dB value that was used in the active-learning process. However, the standard deviations for ears 1R, 8L, and 8R were markedly higher.

An important question for active-learning tests is the value of N needed to yield a reliable result. Figure 6 shows three measures of the accuracy of the fit as a function of N. The left panel shows the negative (natural) log probability of predicting the responses for all 100 trials
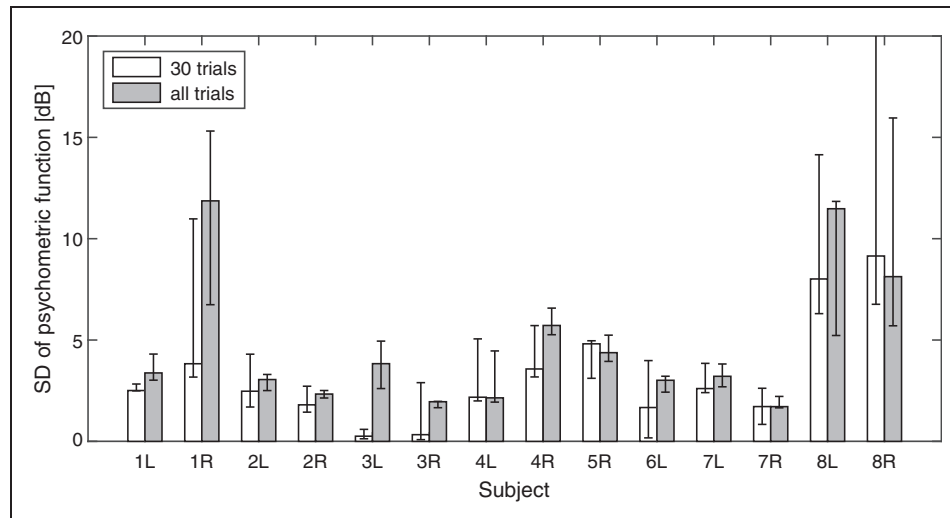
Figure 5. Estimated standard deviation of the psychometric function (a Gaussian CDF) for each ear after 30 trials (white bars) and 100 trials (gray bars). Bars show the median of the three runs, and error bars show the range.
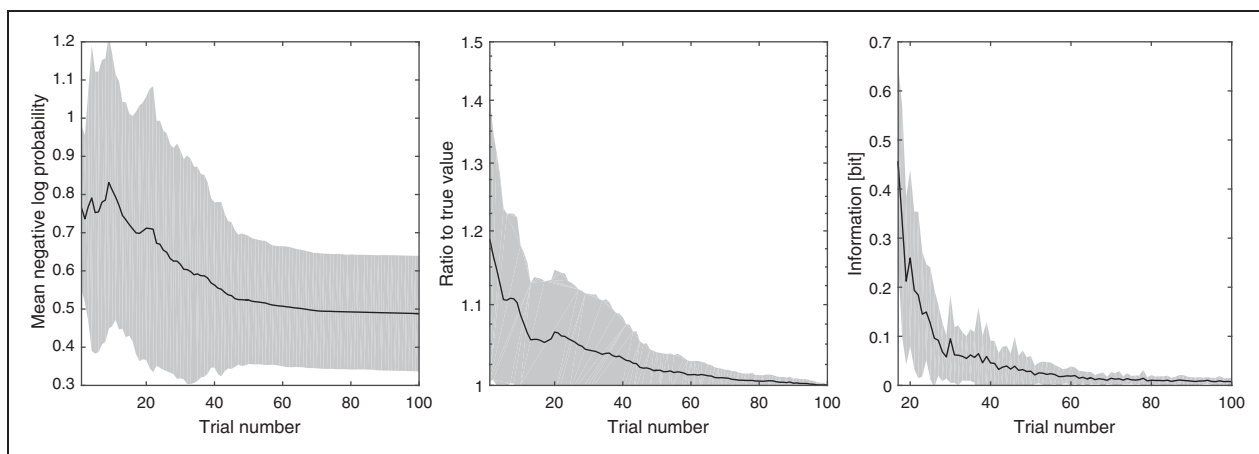
Figure 6. Mean negative (natural) log probability of predicting responses for all 100 trials correctly from the responses for the first N trials (left), ratio of $\hat{f}_{e,n \text{ trials}}/\hat{f}_{e,100 \text{ trials}}$ or its reciprocal if smaller than 1 (middle), and the queried information (right) as a function of trial number. Solid lines show the mean across ears and runs, and gray areas show ± 1 standard deviation. All three measures are close to asymptotic values after about 50 trials.

correctly from the responses for the first N trials, divided by 100. The lower this number, the more accurate are the predictions. The middle panel shows the ratio between the most likely value of f_e after N trials and the most likely value after 100 trials, or its reciprocal if the ratio was smaller than 1. The right panel shows the mutual information that was queried in the Nth trial, which declines from a theoretical maximum of 1 bit to 0.1 bit after about 25 trials, before reaching an asymptote of 0. For all panels, solid lines show the means across ears and runs, and gray areas show ± 1 standard deviation. All measures are close to asymptotic values after about 50 trials. The mutual information (right panel) is available during a run, that is, it can be calculated with the

knowledge that is available after the Nth trial. It is highly correlated with the mean negative log probability, $r(98) = 0.88$, $p < .001$, and the ratio of the estimate of f_e to the true value, $r(98) = 0.84$, $p < .001$. Hence, the mutual information could be used to decide when f_e was determined with sufficient precision for a run to be terminated.

Discussion

As shown in the bottom-right panel of Figure 1, a basal DR could start at a frequency where the audiometric threshold was only slightly higher than normal. More generally, the audiometric threshold at f_e varied widely

across ears. Also, the slope of the audiogram for frequencies close to f_e varied widely across test ears. This is consistent with previous results showing that the presence and edge frequency of a DR cannot be diagnosed reliably from the audiogram (Aazh & Moore, 2007; Vinay & Moore, 2007).

The open symbols in Figure 2 show that the estimates of f_e from the three Smart DRT runs were close to each other, that is, the active-learning procedure led to reproducible results. This was the case even when the Fast PTCs failed to provide a clear result, although for the subjects for whom this was the case the Smart DRT results varied more across runs than for the other subjects. The estimates of f_e obtained using the Smart DRT were close to the estimates based on "successful" Fast PTCs, and within the range of plausible values of f_{tip} given the difficulty of estimating f_{tip} from the irregular PTCs. The similarity of the estimates of f_e for the Smart DRT and the Fast PTC methods suggests that the values estimated using the Smart DRT are close to the "true" values.

The estimates of f_e obtained by Salorio-Corbetto et al. (2017) on the basis of several Fast PTCs agreed with two of the three Smart DRT estimates in cases where the present Fast PTCs were not successful. Altogether, the results of the Smart DRT were consistent with estimates obtained using other procedures in 13 of 14 cases. For the remaining ear (8R), the Smart DRT runs indicated a need to be cautious in interpreting the results: Responses to the initial grid were markedly different across runs (not shown in the results section), suggesting that he was not consistent in his criterion for making a decision.

In the Smart DRT, the values of f_{mask} and L_{mask} for the next trial were chosen to provide information about f_e and OHCL(f_e) but with some specific constraints to ensure exploration of a reasonable range of the parameter space. In theory, this is somewhat less efficient than choosing the parameters that provide the maximum information, but it prevented the model-based approach from becoming "stuck" in an inappropriate region of the parameter space and it ensured some randomness in the stimulus parameters for successive trials, preventing the model from overestimating the precision of its estimates of the model parameters.

The consistency of the performance of the subjects was determined using two measures, the false alarm rate and the steepness of the psychometric function. The two subjects who produced false positives at a rate of 10% or more both reported having tinnitus, and this may have contributed to their relatively high false-positive rates, if the pitch of their tinnitus was close to that of the signal (see also Lentz, Walker, Short, & Skinner, 2017). For the other subjects, the false alarm rate of 1% on average was exactly the fixed value that was applied in Equation 1. Thus, the active-learning

procedure used reasonable assumptions for most subjects but was overconfident for the two outliers. In practice, it is difficult to adjust the false alarm rate to the value appropriate for an individual, since many trials are needed to achieve a reasonable estimate. Nevertheless interleaved catch trials can be used as a check to be evaluated at the end of a test.

For 11 of the 14 ears, the standard deviation of the Gaussian CDF used to estimate the psychometric function was reasonably close to the value of 3 dB that was assumed in Equation 1 (mean: 3.3 dB, range: 1.9–5.8 dB). However, the standard deviation of the psychometric function was considerably higher, at about 10 dB, for 1R, 8L, and 8R. The high value for 1R may have been a consequence of tinnitus. Nonetheless the final estimate of f_e for 1R was close to that obtained from the Fast PTCs and that obtained by Salorio-Corbetto et al. (2017). This suggests that the Smart DRT worked for him, but more trials were needed than for the other subjects to obtain the same accuracy. In order for the estimate of f_e to fall and stay within 0.3 Cams of the final estimate, 63, 50, and 45 trials were needed in his three runs, compared with the average of 33 trials. For 8L and 8R, the relatively flat psychometric functions may reflect his difficulty in distinguishing the tone from the noise. Some subjects with DRs hear a tone whose frequency falls within a DR as being noise-like (Huss & Moore, 2005). Cognitive factors are unlikely to account for the relatively flat psychometric functions, because his psychometric function for the audiogram was steeper than normal, showing a near-perfect separability of yes and no responses.

Ideally, the steepness of the psychometric function should be set to the individual value in Equation 1. However, its optimization is not an explicit aim of the Smart DRT and is not required to obtain a reasonable result. During early trials, perfect separability between yes and no responses may be achieved due to the data being sparsely spaced on the (f_{mask}, L_{mask}) initial grid. After 30 trials, this was the case for both ears of Subject 3 (3L and 3R), and at least one run for 6L (see Figure 5). Underestimation of the standard deviation of the psychometric function leads to overconfidence in the current estimates of f_e and OHCL(f_e). This can have two disadvantages. First, it may cause the procedure to stick to wrong estimates by preventing it from querying parameter values around the true values. Second, overconfidence leads to an underestimation of variance or uncertainty, which may be used as a criterion for stopping the test. A suitable strategy for optimizing the steepness of the psychometric function would be as follows. For the first few active-learning trials, set the standard deviation of the CDF to 3 dB. Afterwards, set the standard deviation to the estimated value with the restriction that it is at least 2 dB. If it is bigger than 20 dB, the run should be restarted.

The number of trials should be chosen to provide an estimate that is accurate enough for fitting a hearing aid. It has been recommended that amplification should not be provided for frequencies higher than about $1.7f_e$ (Baer, Moore, & Kluk, 2002). A precision of 0.3 Cams corresponds to a factor of less than 1.05 in frequency for frequencies between 500 and 4000 Hz. When this precision is reached, the cutoff frequency used for fitting would be between 1.6 and $1.8f_e$, which is accurate enough for an initial fitting. A ratio of about 1.05 or better was reached on average after about 30 trials (Figure 6, middle). The ratio was below 1.05 after 50 trials even for the worst cases. Using more than this number of trials did not increase the predictability of the entire data set (Figure 6, left). Fifty trials of the Smart DRT plus some catch trials and one run of the Quick TEN(HL) test were typically done in 5 to 8 min. The test could be even faster for some subjects if the stopping criterion was based on the information obtained in the next trial (Figure 6, right) or the uncertainty about f_e (see Figure 2). If this approach is adopted, the individual consistency of responses (Figure 5) should be taken into account, since it directly affects the uncertainty about f_e.

Conclusions

The present study proposed and evaluated an active-learning method, Smart DRT, for quickly detecting a basal DR and estimating its edge frequency, f_e. The duration required for testing one ear was 5 to 8 min. The estimates of f_e were close to those obtained using other methods for 13 of 14 test ears, which is a slightly higher success rate than for Fast PTCs when the same amount of time was spent. The Smart DRT may be useful when fitting a hearing aid, providing accurate estimates of f_e in a short time.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Aazh, H., & Moore, B. C. J. (2007). Dead regions in the cochlea at 4 kHz in elderly adults: Relation to absolute threshold, steepness of audiogram, and pure-tone average. *Journal of the American Academy of Audiology*, *18*, 97–106. doi:10.3766/jaaa.18.2.2

Alcántara, J. I., & Moore, B. C. J. (2002). The relative role of beats and combination tones in determining the shapes of masking patterns: II. Hearing-impaired listeners. *Hearing Research*, *165*, 103–116. doi:10.1016/S0378-5955(02)00291-5

Baer, T., Moore, B. C. J., & Kluk, K. (2002). Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *Journal of the Acoustical Society of America*, *112*, 1133–1144. doi:10.1121/1.1498853

Borg, E., Canlon, B., & Engström, B. (1995). Noise-induced hearing loss. Literature review and experiments in rabbits. Morphological and electrophysiological features, exposure parameters and temporal factors, variability and interactions. *Scandinavian Audiology Supplementum*, *40*, 1.

British Society of Audiology. (2011). *Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking: Recommended procedure*. Reading, England: British Society of Audiology.

Burkhard, M. D., & Sachs, R. M. (1975). Anthropometric manikin for acoustic research. *Journal of the Acoustical Society of America*, *58*, 214–222. doi:10.1121/1.380648

Chistovich, L. (1957). Frequency characteristics of masking effect. *Biofizika*, *2*, 714–725.

DiMattina, C. (2015). Fast adaptive estimation of multidimensional psychometric functions. *Journal of Vision*, *15*, 1–20. doi:10.1167/15.9.5

Engström, B. (1983). Stereocilia of sensory cells in normal and hearing impaired ears: A morphological, physiological and behavioural study. *Scandinavian Audiology*, *19*, 1–34.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*, 103–138. doi:10.1016/0378-5955(90)90170-T

Green, D. M., Birdsall, T. G., & Tanner, W. P., Jr. (1957). Signal detection as a function of signal intensity and duration. *The Journal of the Acoustical Society of America*, *29*, 523–531. doi:10.1121/1.1908951

Green, D. M., & Sweets, J. A. (1974). *Signal detection theory and psychophysics*. New York, NY: Krieger.

Houlsby, N. M. T., Huszár, F., Ghahramani, Z., & Lengyel, M. (2011). Bayesian active learning for classification and preference learning. *arXiv Preprint*, arXiv:1112.5745v1.

Houlsby, N. M. T., Huszár, F., Ghassemi, M. M., Orbán, G., Wolpert, D. M., & Lengyel, M. (2013). Cognitive tomography reveals complex, task-independent mental representations. *Current Biology*, *23*, 2169–2175. doi:10.1016/j.cub.2013.09.012

Huss, M., & Moore, B. C. J. (2005). Dead regions and noisiness of pure tones. *International Journal of Audiology*, *44*, 599–611. doi:10.1080/02640410500243962

Kluk, K., & Moore, B. C. J. (2004). Factors affecting psychophysical tuning curves for normally hearing subjects. *Hearing Research*, *194*, 118–134. doi:10.1016/j.heares.2004.04.012

Kluk, K., & Moore, B. C. J. (2005). Factors affecting psycho-physical tuning curves for hearing-impaired subjects with high-frequency dead regions. *Hearing Research*, *200*, 115–131. doi:10.1016/j.heares.2004.09.003

Kluk, K., & Moore, B. C. J. (2006). Detecting dead regions using psychophysical tuning curves: A comparison of simultaneous and forward masking. *International Journal of Audiology*, *45*, 463–476. doi:10.1080/14992020600753189

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*, 2729–2737. doi:10.1016/S0042-6989(98)00285-5

Kujala, J. V., & Lukka, T. J. (2006). Bayesian adaptive estimation: The next dimension. *Journal of Mathematical Psychology*, *50*, 369–389. doi:10.1016/j.jmp.2005.12.005

Lentz, J. J., Walker, M. A., Short, C. E., & Skinner, K. G. (2017). Audiometric testing with pulsed, steady, and warble tones in listeners with tinnitus and hearing loss. *American Journal of Audiology*, *26*, 328–337. doi:10.1044/2017_AJA-17-0009

Moore, B. C. J. (2001). Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids. *Trends in Amplification*, *5*, 1–34. doi:10.1177/108471380100500102

Moore, B. C. J. (2004). Dead regions in the cochlea: Conceptual foundations, diagnosis, and clinical applications. *Ear and Hearing*, *25*, 98–116. doi:10.1097/01.aud.0000120359.49711.d7

Moore, B. C. J. (2007). *Cochlear hearing loss: Physiological, psychological and technical issues*. Chichester, England: John Wiley & Sons.

Moore, B. C. J. (2012). *An introduction to the psychology of hearing*. Leiden, The Netherlands: Brill.

Moore, B. C. J., & Alcántara, J. I. (2001). The use of psycho-physical tuning curves to explore dead regions in the cochlea. *Ear and Hearing*, *22*, 268–278. doi:10.1097/00003446-200108000-00002

Moore, B. C. J., & Glasberg, B. R. (2004). A revised model of loudness perception applied to cochlear hearing loss. *Hearing Research*, *180*, 70–88. doi:10.1016/S0378-5955(03)00347-2

Moore, B. C. J., Glasberg, B., & Schlueter, A. (2010). Detection of dead regions in the cochlea: Relevance for combined electric and acoustic stimulation. In P. van de Heyning, & A. Kleine Punte (Eds.), *Cochlear implants and hearing preservation* (pp. 43–50). Basel, Switzerland: Karger.

Moore, B. C. J., Glasberg, B. R., & Stone, M. A. (2004). New version of the TEN test with calibrations in dB HL. *Ear and Hearing*, *25*, 478–487. doi:10.1097/01.aud.0000145992.31135.89

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., & Alcántara, J. I. (2000). A test for the diagnosis of dead regions in the cochlea. *British Journal of Audiology*, *34*, 205–224. doi:10.3109/03005364000000131

Moore, B. C. J., & Malicka, A. N. (2013). Cochlear dead regions in adults and children: Diagnosis and clinical implications. *Seminars in Hearing*, *34*, 37–50.

Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian processes for machine learning*. Cambridge, MA: MIT Press.

Salorio-Corbetto, M., Baer, T., & Moore, B. C. J. (2017). Quality ratings of frequency-compressed speech by participants with extensive high-frequency dead regions in the cochlea. *International Journal of Audiology*, *56*, 106–120. doi:10.1080/14992027.2016.1234071

Schuknecht, H. F., & Woellner, R. C. (1953). Hearing losses following partial section of the cochlear nerve. *The Laryngoscope*, *63*, 441–465. doi:10.1288/00005537-195306000-00001

Sęk, A., Alcántara, J. I., Moore, B. C. J., Kluk, K., & Wicher, A. (2005). Development of a fast method for determining psychophysical tuning curves. *International Journal of Audiology*, *44*, 408–420. doi:10.1080/14992020500060800

Sęk, A., & Moore, B. C. J. (2011). Implementation of a fast method for measuring psychophysical tuning curves. *International Journal of Audiology*, *50*, 237–242. doi:10.3109/14992027.2010.550636

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423, 623–656.

Shen, Y., & Richards, V. M. (2013). Bayesian adaptive estimation of the auditory filter. *Journal of the Acoustical Society of America*, *134*, 1134–1145. doi:10.1121/1.4812856

Song, X. D., Sukesan, K. A., & Barbour, D. L. (2018). Bayesian active probabilistic classification for psychometric field estimation. *Attention, Perception, & Psychophysics*, *80*, 798–812. doi:10.3758/s13414-017-1460-0

Song, X. D., Wallace, B. M., Gardner, J. R., Ledbetter, N. M., Weinberger, K. Q., & Barbour, D. L. (2015). Fast, continuous audiogram estimation using machine learning. *Ear and Hearing*, *36*, e326–e335. doi:10.1097/AUD.0000000000000186

Vinay, & Moore, B. C. J. (2007). Prevalence of dead regions in subjects with sensorineural hearing loss. *Ear and Hearing*, *28*, 231–241. doi:10.1097/AUD.0b013e31803126e2

Zhang, T., Dorman, M. F., Gifford, R., & Moore, B. C. J. (2014). Cochlear dead regions constrain the benefit of combining acoustic stimulation with electric stimulation. *Ear and Hearing*, *35*, 410. doi:10.1097/AUD.0000000000000032