Research Article

Toward Routine Assessments of Auditory Filter Shape

Yi Shen,^a Allison B. Kern,^a and Virginia M. Richards^b

Purpose: A Bayesian adaptive procedure, that is, the quick auditory filter (qAF) procedure, has been shown to improve the efficiency for estimating auditory filter shapes of listeners with normal hearing. The current study evaluates the accuracy and test–retest reliability of the qAF procedure for naïve listeners with a variety of ages and hearing status.

Method: Fifty listeners who were naïve to psychophysical experiments and exhibit wide ranges of age (19–70 years) and hearing threshold (–5 to 70 dB HL at 2 kHz) were recruited. Their auditory filter shapes were estimated for a 15-dB SL target tone at 2 kHz using both the qAF procedure and the traditional threshold-based procedure. The auditory filter model was defined using 3 parameters: (a) the sharpness of the tip portion of the auditory filter, p ; (b) the prominence of the low-frequency tail of the filter,

$10\log(w)$; and (c) the listener’s efficiency in detection, $10\log(K)$.

Results: The estimated parameters of the auditory filter model were consistent between 2 qAF runs tested on 2 separate days. The parameter estimates from the 2 qAF runs also agreed well with those estimated using the traditional procedure despite being substantially faster. Across the 3 auditory filter estimates, the dependence of the auditory filter parameters on listener age and hearing threshold was consistent across procedures, as well as consistent with previously published estimates.

Conclusions: The qAF procedure demonstrates satisfactory test–retest reliability and good agreement to the traditional procedure for listeners with a wide range of ages and with hearing status ranging from normal hearing to moderate hearing impairment.

Frequency selectivity is a fundamental property of auditory perception. It allows the encoding of spectral details, which is essential for signal detection in noise, sound discrimination, pitch perception, and speech understanding. The predominant mechanism underlying frequency selectivity is the filtering of acoustic inputs in the cochlea, where locations near the base of the cochlea respond selectively to high-frequency sounds whereas locations near the apex respond to low-frequency sounds. The sharpness of frequency tuning at each of the cochlear locations is related to the health of the sensory hair cells, especially the outer hair cells (e.g., Moore, 1986; Moore, Vickers, Plack, & Oxenham, 1999). For listeners with sensorineural hearing loss, frequency tuning in the cochlea broadens due to the dysfunction of the outer hair cells. As a consequence, listeners with hearing impairment (HI) exhibit

degradations in psychophysical tasks that require fine frequency selectivity, such as detection of tonal targets in noise and discrimination of spectral profiles (e.g., Moore, 1995). Accordingly, such psychophysical tasks can be developed into noninvasive behavioral tests that probe the integrity of peripheral processing for listeners with HI, which is especially meaningful for clinical practice and clinical research in audiology. Recently, various time-efficient behavioral assessment tools for frequency selectivity have been proposed. The current study aims to validate one of these psychophysical procedures based on Bayesian adaptive estimation, that is, the quick auditory filter (qAF) procedure (Shen & Richards, 2013a; Shen, Sivakumar, & Richards, 2014), and to evaluate the test–retest reliability of the procedure using a relatively large number of listeners with wide ranges of ages and hearing status.

Over decades of research, the behavioral estimation of frequency selectivity of the auditory system has evolved into a few well-established psychophysical paradigms, including, for example, the measurement of the psychophysical tuning curve (PTC; e.g., Vogten, 1978) and the auditory filter shape (Glasberg & Moore, 1990; Patterson, 1976). Broader PTCs or auditory filters are indicative of degraded frequency selectivity. These two metrics have

^aDepartment of Speech and Hearing Sciences, Indiana University Bloomington

^bDepartment of Cognitive Sciences, University of California, Irvine

Correspondence to Yi Shen: shen2@indiana.edu

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Antje Ihlefeld

Received March 16, 2018

Revision received June 26, 2018

Accepted October 11, 2018

https://doi.org/10.1044/2018_JSLHR-H-18-0092

Disclosure: The authors have declared that no competing interests existed at the time of publication.

been repeatedly used in clinical studies that evaluated the psychoacoustical basis of speech recognition (e.g., measurements of PTC in Arlinger & Dryselius, 1990; Stelmachowicz, Jesteadt, Gorga, & Mott, 1985; measurements of the auditory filter shape in Badri, Siegel, & Wright, 2011; Davies-Venn, Nelson, & Souza, 2015; Dubno & Dirks, 1989; Dubno & Schaefer, 1992; Hopkins & Moore, 2011; Jin & Nelson, 2010; Kortlang, Mauermann, & Ewert, 2016; Leek & Summers, 1996; Neher, Laugesen, Jensen, & Kragelund, 2011; Strelcyk & Dau, 2009; Summers, Makashay, Theodoroff, & Leek, 2013).

Despite being widely adopted in clinical research, the standard estimation procedures for the PTC and auditory filter shape require time-consuming data collection, often leading to hours of testing time. Consequently, research efforts have been made to develop time-efficient assessments of frequency selectivity. One of the approaches utilizes Békésy tracking during the PTC measurement, preventing the need of repetitively estimating detection thresholds (e.g., Charaziak, Souza, & Siegel, 2012; Malicka, Munro, & Baker, 2009; Myers & Malicka, 2014; Pepler, Munro, Lewis, & Kluk, 2014; Şek, Alcántara, Moore, Kluk, & Wicher, 2005). An alternative approach is to maintain the classical notched-noise masking paradigm while optimizing stimulus sampling during the measurement of the auditory filter shape (e.g., Shen & Richards, 2013a; Shen et al., 2014).

Traditionally, the auditory filter shape is estimated using a threshold-based approach. That is, the detectability of a tonal target is measured in the presence of a masker that is composed of two noise bands, one above and one below the target frequency, forming a spectral notch. The detection threshold of a tone at the target frequency is estimated for various notch widths, with the notch being either symmetric or asymmetric about the target frequency. The estimated thresholds are then used to fit the auditory filter model (e.g., Baker & Rosen, 2002, 2006; Glasberg & Moore, 1990; Patterson, Nimmo-Smith, Weber, & Milroy, 1982; Rosen, Baker, & Darling, 1998). The shape of the auditory filter can be formulated in various ways, but a commonly adopted one uses the rounded-exponential (roex) functions (Patterson et al., 1982). Figure 1 shows two examples of the auditory filter shapes formulated using the roex functions for listeners with normal hearing (NH) at 2 kHz (from Baker & Rosen, 2006). The roex auditory filter is a band-pass filter with a sharp tip. The auditory filter shape is also asymmetric, especially at high levels, where the degree of filter asymmetry is governed by the prominence of the shallow low-frequency tail (e.g., Baker & Rosen, 2006; Shen et al., 2014).

To estimate the auditory filter shape, notch widths are often sampled in fine steps to capture the sharp filter tip and the transition from the tip to the tail of the filter. Moreover, asymmetric notch placements are used to capture the asymmetry of the filter shape. This requires a large number of notch conditions, and so data collection is quite time-consuming. A typical experiment estimating the auditory filter shape at one frequency takes more than 2,000 experimental trials and up to 3 hr of testing. Previous studies

Figure 1. The auditory filter shape at 2 kHz estimated by Baker and Rosen (2006) for target levels of 30 dB SPL (left) and 50 dB SPL (right). The filter shapes shown are formulated using the rounded-exponential (roex) function. The key parameters of the filter shape, that is, the tip sharpness p and the tail prominence $10\log w$, are illustrated.

have suggested that certain notch conditions may be omitted to improve the efficiency of data collection (Leeuw & Dreschler, 1994; Stone, Glasberg, & Moore, 1992). However, the notch conditions that can be removed without significantly altering the filter shape estimate may be different for different listeners, especially for the population with HI. A shortened procedure with a preselected subset of notch conditions is not likely to be flexible enough to adapt to individual listeners' auditory filter shapes.

In contrast to the traditional threshold-based approach, the Bayesian adaptive procedure proposed by Shen and Richards (2013a) and Shen et al. (2014), namely, the qAF procedure, adopts a parameter-based approach. The qAF procedure is among the emerging efforts to introduce Bayesian adaptive testing for the estimation of high-dimensional psychophysical functions or models. Besides the auditory filter, Bayesian adaptive procedures have also been utilized to assess the audiometric thresholds (Song, Garnett, & Barbour, 2017; Song et al., 2015), the temporal modulation transfer function (Shen & Richards, 2013b), the equal-loudness contour (Shen, Zhang, & Zhang, 2018), and the speech intelligibility index (Shen & Kern, 2018). In the domain of vision research, there are also established Bayesian adaptive procedures for the assessments of visual function (e.g., Dorr et al., 2017; Lesmes, Lu, Baek, & Albright, 2010). Recently, a general framework for implementing such Bayesian adaptive procedures for customizable applications, that is, the QUEST+ procedure, has been proposed (Watson, 2017). The main premise of the qAF procedure is that it iteratively optimizes stimulus selection following each trial to achieve direct estimation of the parameters for the auditory filter shape of interest without estimating detection thresholds first.

Shen and Richards (2013a) demonstrated that the qAF procedure provided estimates of the auditory filter shape that well-predicted detection thresholds collected using the traditional threshold-based approach. Shen et al. (2014) further developed the computational algorithms associated with the qAF procedure to enable the estimation

of high-parameter models. This study also evaluated the qAF procedure under both simultaneous and forward masking conditions using a group of listeners who were naïve to psychoacoustic experiments. However, these previous investigations of the qAF procedure were limited in that (a) only young listeners with NH were tested and (b) the individual differences in the estimated auditory filter shapes were not studied due to the relatively small sample sizes.

Extending the validation of the qAF procedure to a more diverse population is necessary because previous studies of the effect of hearing threshold on the auditory filter shape have shown that the individual differences in the filter shape typically increase as the degree of hearing loss increases (e.g., Dubno & Dirks, 1989; Glasberg & Moore, 1986; Laroche, Hétu, Quoc, Josseland, & Glasberg, 1992; Sommers & Humes, 1993). Therefore, it is possible that the satisfactory test-retest reliability demonstrated using listeners with NH may not hold when listeners with HI are tested. The current study evaluates the reliability of the qAF procedure using a large group of naïve listeners with wide ranges of ages and hearing status. Moreover, the current study also measures auditory filter shapes using the traditional threshold-based procedure to evaluate the agreement between the qAF and threshold-based procedures.

Method

Listeners

Fifty adult listeners (25 women and 25 men, between 18 and 78 years) were recruited, all of whom were naïve to psychoacoustic experiments. The subject recruitment did not exclude any listeners on the basis of age or hearing ability. All listeners provided informed consent before participating. The experimental protocol for this study was approved by the institutional review board at Indiana University. Hearing thresholds were measured from all listeners using standardized audiometric procedures. The ear with the lower pure-tone average (PTA) threshold (i.e., the mean of the hearing thresholds at 0.5, 1, and 2 kHz) was tested. If the PTA was the same in both ears, the left ear was tested.

Figure 2 plots the audiometric threshold as a function of age for the listeners in the current study. The listeners ranged broadly in both age and hearing threshold. The distribution of listeners' ages was somewhat bimodal with most listeners either below 30 or above 50 years. Most listeners had either NH or slight hearing loss; only a small subset of the listeners had mild or moderate hearing loss. A significant correlation between age and PTA threshold was observed ($r = .54, p < .001$).

Stimuli

Listeners were instructed to detect the presence of a pure-tone target in simultaneous noise. Three sound intervals were presented on each trial, separated by 400-ms interstimulus intervals. In each interval, a masker sound was presented, which contained two noise bands, one on

either side of the frequency of the target tone (f_t). The masker duration was 300 ms, including 10-ms onset and offset raised cosine ramps. In one of the three intervals, determined randomly for each trial, the target tone was presented and gated on and off simultaneously with the masker. During data collection, the level of the target tone was fixed at 15 dB SL. The level for 0 dB SL was defined by the hearing threshold measured at the target frequency using the standard audiometric procedure.

The two masker bands, each of which had a bandwidth of $0.25f_t$, formed a spectral notch around the target. The notch width of the two bands, that is, the distance between the inner edge of the noise bands and f_t , was expressed as the normalized distance $g = |f - f_t| / f_t$, where f is frequency in Hz. Depending on the experimental condition, the notch width on the low-frequency side (lower notch width g_l) might be equal to or different from the notch width on the high-frequency side (upper notch width g_h) of the target frequency, forming symmetric or asymmetric notches, respectively. The spectrum level of the masker L_M was adaptively adjusted on every trial (see Rosen & Baker, 1994, for a discussion on the potential advantages of varying the masker level over varying the target level in auditory filter measurements).

All stimuli were digitally generated and presented at a sampling rate of 22050 Hz. They were presented to the listeners via a 24-bit sound card (Microbook II, Mark of the Unicorn, Inc.) and headphones (HD280 Pro, Sennheiser Electronic GmbH & Co. KG). Six of the listeners were tested in a quiet office, and the remaining listeners were tested in sound-attenuating booths.

Procedure

General Procedure

Audiometric thresholds were first measured for each listener, followed by an estimate of the auditory filter shape at 2 kHz using the qAF procedure (i.e., the first 2-kHz qAF run). Although no training was provided before the first 2-kHz qAF run, the experimental task was explained to the listener verbally. The experiment began when the listener confirmed that he or she understood the instruction. After the first 2-kHz qAF run, the auditory filter shape at 2 kHz was estimated again using the threshold-based procedure. Finally, the auditory filter shape at 2 kHz was estimated using the qAF procedure a second time (i.e., the second 2-kHz qAF run). Each run of the qAF procedure took approximately 15 min to complete, whereas the threshold-based procedure took about 120 min. The experiment was completed in two 2-hr experimental sessions, and the first and second 2-kHz qAF runs were always tested on different days. For a subset of the listeners (31 of 50 listeners), additional testing time was available at the end of the second test session, so the auditory filter shape at 1 kHz was also estimated (i.e., the 1-kHz qAF run). These data provided an opportunity to investigate the effects of listener age and hearing threshold on the auditory filter shape at a target frequency other than 2 kHz. The target frequencies of 1 and 2 kHz were chosen in the current study because these

Figure 2. The scatter plot of audiometric threshold at 2 kHz (left) and 1 kHz (right) as a function of age for all listeners in the current study.

frequency regions are known to be important for speech understanding (American National Standards Institute, 1997).

The first 2-kHz qAF run reflects results expected from naïve listeners, and the test–retest reliability of the qAF procedure was provided by comparing the estimated parameters of the auditory filter model from the first and second 2-kHz qAF runs. The agreement between the qAF and threshold-based procedures was evaluated by comparing the first 2-kHz qAF run to the threshold-based procedure.

In the current study, the auditory filter shape was formulated as a bandpass filter centered at the target frequency. The shape of the filter was defined by two parameters (p and $10\log w$; see Figure 1). The dimensionless parameter p indicates the sharpness of the tip portion of the auditory filter, and the parameter $10\log w$ describes the contribution of the tail portion to the total filter output in decibels. An additional model parameter, $10\log K$, was included to indicate detection efficiency, with higher values of $10\log K$ associated with higher signal-to-noise ratio required to reach detection threshold. Each estimate of the auditory filter shape, using either the qAF or threshold-based procedure, included the estimates of the three model parameters (i.e., tip sharpness p , tail prominence $10\log w$, and detection efficiency $10\log K$). The difference between the qAF and threshold-based procedures was that the qAF procedure determined all three stimulus parameters (masker spectrum level L_M , upper notch width g_h , and lower notch width g_l) adaptively on a trial-by-trial basis whereas the threshold-based procedure tested the predetermined notch conditions (upper notch width g_h and lower notch width g_l) in blocks. The implementation details of the qAF and threshold-based procedures are provided below for interested readers. For the qAF procedure, a mathematical description of the adaptive stimulus selection algorithm is provided in the Appendix.

The qAF Procedure

The qAF procedure implemented was similar to that described by Shen et al. (2014), with a few exceptions (the implementation details are provided in the Appendix). At

the beginning of the qAF procedure, a prior distribution was defined for the three model parameters (tip sharpness p , tail prominence $10\log w$, and detection efficiency $10\log K$), which described the “initial guess” of the parameter values. Based on computer simulations, three different prior configurations were used that differed in the prior mean for $10\log w$ and prior variance for $10\log K$. These differences in prior distributions did not yield significant differences in results and so are not addressed further (see the Appendix for details).

After each trial of the qAF procedure, by combining the stimulus parameters and whether a correct response was observed, Bayes’ theorem was used to update the estimates of the parameters p , $10\log w$, and $10\log K$. The updated parameter estimates were then used to optimize the choice of stimulus parameters to be used on the following trial. An entropy-based criterion was used during this stimulus optimization process so that the expected information gain from the next trial was maximized (e.g., Kontsevich & Tyler, 1999; Lesmes, Jeon, Lu, & Doshier, 2006; Lesmes et al., 2010; Shen & Kern, 2018; Shen & Richards, 2013a, 2013b; Shen et al., 2014).

Each qAF run consisted of four experimental blocks of 50 trials each. Each block began with a warm-up phase (included in the 50 trials) followed by the qAF trials. During the warm-up phase, the masker was configured as the masker that would be presented on the first qAF trial of the block. The target level during the warm-up phase was adaptively varied using a two-down, one-up staircase procedure (Levitt, 1971). The initial target level was 45 dB SL. For listeners with very high thresholds at the target frequency, a target level of 45 dB SL corresponded to a sound pressure level that would exceed the safety limit. For these cases, the initial target level was set at 100 dB SPL. The target level was reduced by 10 dB, following two consecutive correct responses, and increased by 10 dB, following a single incorrect response. The warm-up phase terminated once two reversals were reached. Then, the qAF trials began, the target level was fixed at 15 dB SL, and the masker characteristics

(defined by masker spectrum level L_M , upper notch width g_h , and lower notch width g_l) were varied adaptively. For the second through the fourth block in a qAF run, the qAF trials resumed from the last qAF trial of the previous block. This led to approximately a total of 40 warm-up trials and 160 qAF trials per qAF run.

The Threshold-Based Procedure

The threshold-based procedure adopted in the current study followed that described by Baker and Rosen (2006). Alternative configurations of the threshold-based procedure are available in the literature, and some of these configurations have been designed to improve the time efficiency of the procedure by omitting certain less informative test conditions (e.g., Stone et al., 1992). Because the focus in the current study was not to explicitly compare the time efficiencies between the qAF and threshold-based procedures, a typical configuration with a relatively complete set of test conditions was selected.

The masker levels at the detection threshold of the target tone were estimated for various notch configurations. These notch configurations included six symmetric notch conditions, where both g_h and g_l were set to 0, 0.1, 0.2, 0.3, 0.4, or 0.5, and 10 asymmetric notch conditions. In five of the asymmetric notch conditions, g_h was smaller than g_l by 0.2 and g_h was 0, 0.1, 0.2, 0.3, or 0.4. In the other five asymmetric notch conditions, g_l was smaller than g_h by 0.2 and g_l was 0, 0.1, 0.2, 0.3, or 0.4. The masker levels at threshold were estimated for all 16 notch configurations in random order and repeated once with an independent random sequence. After data collection, the average thresholds from the two estimates were used to fit the auditory filter model. The form of the auditory filter model was the same as described for the qAF procedure.

To estimate each masker level at threshold, the masker spectrum level L_M was adaptively varied using a one-down, two-up staircase procedure (Levitt, 1971). The initial masker spectrum level was 0 dB SPL. The masker spectrum level was increased following two consecutive correct responses and decreased following a single incorrect response. The initial step size was 10 dB. The step size was reduced to 5 dB after two reversals of the adaptive track and was reduced further to 2 dB after the fourth reversal. The track terminated when 10 reversals were reached, and the masker level at threshold was the average of the masker spectrum level at the last six reversals. The masker levels at threshold under various notch conditions were used to fit the auditory filter model (Equation A5). The estimates of tip sharpness p , tail prominence $10\log w$, and detection efficiency $10\log K$ were obtained using the MATLAB function *fminsearch*, with the initial search points set to 30, -10, and 0, respectively. The fitting process was constrained so that the estimated parameters were always within a set of specified limits (i.e., $20 \leq p \leq 70$, $-60 \text{ dB} \leq 10\log w \leq 0 \text{ dB}$, and $-10 \text{ dB} < 10\log K < 10 \text{ dB}$). The constraints on the ranges of parameter values were implemented to prevent extreme parameter estimates when the model fits were relatively poor. For eight of the 50 listeners, at least one of the parameter estimates reached

the set limits. Nonetheless, the fitted parameters from these listeners were included in the data analyses.

Results

The Test–Retest Reliability of the qAF Procedure

The test–retest reliability of the qAF procedure was evaluated by comparing the estimates of the auditory filter parameters from the first and second 2-kHz qAF runs. As shown in Figure 3, the parameter estimates for tip sharpness p (top left), tail prominence $10\log w$ (top right), and detection efficiency $10\log K$ (bottom left) agreed well between the first and second 2-kHz qAF runs. Although the two qAF runs were conducted on different days, the deviation between the parameter estimates from the two runs is substantially smaller than the individual differences across the listeners.

A closer inspection of Figure 3 suggests that the largest deviations between the two qAF runs is often associated with parameter estimates at the boundaries of the parameter space (occurring for four listeners for the first 2-kHz qAF run and six listeners for the second 2-kHz qAF run). For example, the p estimates (see Figure 3, top left) were occasionally at either 20 or 70. Some of these incidences led to large amounts of deviations between p estimates across the two qAF runs (greater distances from the diagonal line in the panel). Parameter estimates at the boundaries indicated failures of convergence, and the test–retest reliability for parameter p seemed to be most affected by the occasional instabilities of the qAF procedure compared to $10\log w$ and $10\log K$.

The auditory filter bandwidth from each of the two 2-kHz qAF runs and for each subject was quantified using the conventional equivalent rectangular bandwidth (ERB) measure (see Equation A3). The root-mean-square deviation between the ERB estimates from the two runs was 47 Hz on the linear frequency scale or 0.25 octaves on the logarithmic frequency scale. These deviations were an order of magnitude smaller than the range of individual ERB estimates in the experiment (approximately 530 Hz on the linear frequency scale and 2.5 octaves on the logarithmic frequency scale).

A Bland–Altman analysis was conducted to investigate the agreement between the parameter estimates from the two 2-kHz qAF runs. As shown in Figure 4, no systematic bias was observed for tip sharpness p , $t(49) = -0.30$, $p = .768$, tail prominence $10\log w$, $t(49) = -1.68$, $p = .099$, or filter bandwidth ERB, $t(49) = -0.16$, $p = .877$. A small bias was observed for detection efficiency $10\log K$, $t(49) = -2.29$, $p = .027$. This bias means that, on average, listeners required 1.2 dB less signal-to-noise ratio to detect the target tone for the second qAF run. The 95% coefficients of repeatability (i.e., 1.96 times the standard deviation of the differences in parameter estimates) were 18.4, 6.8 dB, 7.2 dB, and 108.2 Hz for tip sharpness p , tail prominence $10\log w$, detection efficiency $10\log K$, and filter bandwidth ERB, respectively.

To assess whether the test–retest reliability of the qAF observed here was satisfactory, two tolerance limits were

Figure 3. The estimates of p (top left), $10\log w$ (top right), $10\log K$ (bottom left), and equivalent rectangular bandwidth (ERB; bottom right) from the first and second 2-kHz quick auditory filter (qAF) runs. The dashed line in each panel indicates equivalent results from the two qAF runs.

Figure 4. The Bland–Alman plot comparing the estimates of p (top left), $10\log w$ (top right), $10\log K$ (bottom left), and equivalent rectangular bandwidth (ERB; bottom right) from the first and second 2-kHz quick auditory filter (qAF) runs. In each panel, the mean difference and the 95% confidence interval are indicated by solid and short dashed horizontal lines (along with values to the right), respectively. In the bottom right panel, the long dashed curve indicates the difference in the ERB that would lead to a 3-dB threshold change for the detection of a 2-kHz pure tone in broadband noise.

defined (see the long dashed curved in Figure 4, bottom right), which are the biases in the ERB estimates corresponding to 3-dB shifts in the model-predicted threshold (in both positive and negative directions) for detecting a 2-kHz pure tone in broadband noise. All observed differences in the ERB estimates were within these two tolerance limits. That is, the variations in the estimated auditory filter shapes from the two qAF runs would not lead to a more than 3-dB difference in the predicted thresholds for tone detection in broadband noise.

The Agreement Between the qAF and Threshold-Based Procedures

Figure 5 plots the parameter estimates obtained from the first 2-kHz qAF run and the threshold-based procedure. The p estimates from the threshold-based procedure were frequently at or near the boundaries of the parameter space (i.e., 20 and 70), which contributed the discrepancies in the p estimates between the qAF and threshold-based procedures. Even so, close agreement was observed for tail prominence $10\log w$ (top right), detection efficiency $10\log K$ (bottom left), and filter bandwidth ERB (bottom right). The root-mean-square deviation for the ERB was 84.9 Hz on the linear frequency scale and 0.42 octaves on a logarithmic frequency scale.

A Bland–Altman analysis was conducted to investigate the agreement between the parameter estimates from

the first 2-kHz qAF run and the threshold-based procedure (see Figure 6). No systematic bias was observed for tip sharpness p , $t(49) = -1.99$, $p = .052$, or filter bandwidth ERB, $t(49) = 1.40$, $p = .169$. A significant bias was found for tail prominence $10\log w$, $t(49) = 3.58$, $p < .001$, indicating the estimated tail prominence was greater by 2.2 dB, on average, for the threshold-based procedure than the first 2-kHz qAF run. A significant bias was also found for detection efficiency $10\log K$, $t(49) = -6.66$, $p < .001$. This means that the listeners, on average, required 3.6 dB less signal-to-noise ratio to detect the target tone during the threshold-based procedure than the qAF procedure. Moreover, the difference in $10\log K$ estimates from the two procedures was negatively correlated with the mean estimates ($r = -.63$, $p < .001$). This means that the listeners who were not very efficient in detecting the target tone tended to exhibit lower $10\log K$ estimates (i.e., greater efficiency) for the threshold-based procedure than the first 2-kHz qAF procedure. The 95% coefficients of repeatability were 27.9, 8.6 dB, 7.5 dB, and 164.9 Hz for tip sharpness p , tail prominence $10\log w$, detection efficiency $10\log K$, and filter bandwidth ERB, respectively.

To assess whether the agreement between the threshold-based and qAF procedures was satisfactory, two tolerance limits associated with 3-dB shifts in threshold for tone detection in broadband noise were defined (as described before). All but two observed differences in the ERB estimates were within these two tolerance limits. That is, for 48 of

Figure 5. The estimates of p (top left), $10\log w$ (top right), $10\log K$ (bottom left), and equivalent rectangular bandwidth (ERB; bottom right) from the first 2-kHz quick auditory filter (qAF) run and the threshold-based procedure from. The dashed line in each panel indicates equivalent results from the two procedures.

Figure 6. The Bland–Alman plot comparing the estimates of p (top left), $10\log w$ (top right), $10\log K$ (bottom left), and equivalent rectangular bandwidth (ERB; bottom right) from the first 2-kHz quick auditory filter (qAF) run and the threshold-based procedure. In each panel, the mean difference and the 95% confidence interval are indicated by solid and short dashed horizontal lines (along with values to the right), respectively. In the bottom right panel, the long dashed curve indicates the difference in the ERB that would lead to a 3-dB threshold change for the detection of a 2-kHz pure tone in broadband noise.

the 50 listeners, the differences in the estimated auditory filter shapes from the threshold-based and qAF procedures would not lead to more than 3-dB changes in the predicted thresholds for tone detection in broadband noise.

Capturing the Effects of Age and Hearing Threshold

The listeners included in the current study exhibited wide ranges of ages and hearing thresholds. Both listener age (e.g., Peters & Moore, 1992; Sommers & Gehr, 1998; Sommers & Humes, 1993) and HI (e.g., Dubno & Dirks, 1989; Glasberg & Moore, 1986; Laroche et al., 1992) could influence the auditory filter shape. Therefore, it is necessary to verify whether the qAF procedure provides consistent accounts for the effect of age and hearing threshold on the auditory filter shape and whether the dependencies of the auditory filter shape on age and hearing threshold captured by the qAF and threshold-based procedures are in agreement.

For each of the three auditory filter estimates at 2 kHz (i.e., the two 2-kHz qAF runs and the threshold-based procedure), multiple linear regressions were repeated for each of the three parameters (i.e., p , $10\log w$, and $10\log K$), treating age and the audiometric threshold at 2 kHz as the two independent variables. For each of the parameters, the unstandardized regression coefficients provided estimates on the

rate of change in the parameter as functions of age and hearing threshold. For example, if the two 2-kHz qAF runs captured similar dependencies of $10\log w$ on hearing threshold, then the unstandardized regression coefficients (i.e., the change in $10\log w$ per unit change in hearing threshold in dB/dB) for the two qAF runs should be comparable.

The unstandardized regression coefficients are summarized in Figure 7. First, consider the results for the 2-kHz target. For all three estimates of tip sharpness p , neither an effect of age nor hearing threshold at 2 kHz was found. There was no significant effect of age on tail prominence $10\log w$, whereas all three estimates of $10\log w$ increased with increasing hearing thresholds at a rate of approximately 0.5 dB/dB. For detection efficiency $10\log K$, all three estimates of $10\log K$ increased with increasing age, indicating less efficient signal detection at later ages. No significant effect of hearing threshold on $10\log K$ was found. Because the regression coefficients from the two 2-kHz qAF runs were consistent with one another, the qAF procedure was sufficiently reliable to capture the dependence of the auditory filter shape on age and HI. Moreover, the effects of age and hearing loss on the auditory filter shape estimated using the qAF and threshold-based procedures were in close agreement.

A parallel multiple linear regression procedure was performed for the 1-kHz qAF runs completed by 31 of the

Figure 7. The unstandardized regression coefficients for linear regression models with age and hearing threshold at the target frequency as the two independent variables and the auditory filter parameters as the dependent variables. Results for the three different auditory filter parameters (p , $10\log w$, and $10\log K$) are arranged into separate panels. In each panel, different bar styles indicate results from the four estimates of the auditory filter in the current study. Error bars indicate 1 SE of the coefficient estimates with the significant coefficients ($p < .05$) indicated by asterisks.

50 listeners. The dependencies of the model parameters on age and hearing threshold were similar to those found for the 2-kHz auditory filter, with notable exceptions. First, tip sharpness p decreased with increasing hearing thresholds at 1 kHz, which was not observed at 2 kHz. Similar negative correlation between hearing threshold and p has been reported in previous studies for the same frequency region (e.g., Dubno & Dirks, 1989; Dubno & Schaefer, 1991; Glasberg & Moore, 1986). Second, $10\log w$ decreased with age, which was not observed for the 2-kHz target.

Discussion

Agreement With Previous Studies

The current study evaluated the reliability of the qAF procedure and its agreement to the threshold-based procedure using a relatively large group of subjects. This was achieved by comparing the auditory filter parameters across repeated qAF runs (to assess reliability) and between the qAF and threshold-based procedures (to assess agreement between the two procedures).

Although the threshold-based procedure implemented in the current study followed previous studies closely, the auditory filter model was sometimes poorly fit. For example, the tip sharpness p was estimated at the boundaries of the parameter space (i.e., either 20 or 70) for eight of the listeners. This may reflect the fact that each estimate of the

masker level at threshold was based only on two adaptive tracks. Under these circumstances, even though the parameter estimates from the qAF and threshold-based procedures were in generally good agreement in this study, it is of interest to compare the current auditory filter estimates with other data sets. Figure 8 plots the individual ERB estimates relative to the target frequency, as functions of the hearing threshold at the target frequency for the current study (crosses). The predicted ERBs from the regression models using coefficients from Figure 7 are shown as solid curve. Because age did not have a significant effect on the auditory filter shape (governed by the parameters p and $10\log w$), a nominal value of 60 years was used as the age variable in generating the predictions. Results from the two 2-kHz qAF runs, the threshold-based procedure, and the 1-kHz qAF runs are plotted in separate panels. In each panel, results from a selection of previous studies that measured the auditory filter shapes at either 1 or 2 kHz are overlaid on the current data.

The data from the current study overlap with the previously reported data, except that the current cohort of listeners mostly had NH to mild hearing loss whereas results for more severe cases of HI are included in the comparison data plotted in Figure 8. Most of these previous results were within the 95% confidence limits of the model prediction (thick dashed curves). Below a hearing loss of approximately 30 dB HL, the ERB was either constant or increased slightly with increasing hearing threshold at the target frequency. This may be because, for listeners with NH and mild hearing loss, the low-frequency tail of the auditory filter did not contribute to the output of the filter significantly (low values of $10\log w$) and the ERB was mainly dominated by the sharpness of the filter's tip (the parameter p). Between 30 and 50 dB HL, the ERB increased rapidly with increasing thresholds (Dubno & Dirks, 1989; Glasberg & Moore, 1986; Sommers & Humes, 1993). For these listeners, the contribution from the low-frequency tail tended to dominate, and the ERB was driven by the parameter $10\log w$. As the value of $10\log w$ increased toward 0 dB, the ERB increased.

For listeners with more than a moderate degree of hearing loss (> 50 dB HL), the regression model predicted that ERBs plateau, owing to the large 0-dB value of $10\log w$. That is, because the parameter w (Equation A2) is restricted to be between 0 and 1, the maximum ERB that the current implementation of the roex auditory filter allowed would be reached when $10\log w$ was 0 dB ($w = 1$). At 2-kHz, this limit in the auditory filter model led to large discrepancies between the model prediction and the data reported by Dubno and Dirks (1989) for four listeners (triangles in the top right corner of each panel associated with the ERB at 2 kHz). In the study of Dubno and Dirks (1989), the auditory filter shape was formulated as the sum of a symmetric tip filter and a flat tail on both high- and low-frequency sides (instead of a shallow low-frequency tail in the current model). This formulation allows the auditory filters to have extremely large ERBs. The ERB estimates for listeners with hearing threshold of greater than 50 dB HL are likely to be sensitive to various formulations of the auditory filter, especially the

Figure 8. Individual equivalent rectangular bandwidth (ERB) estimates (crosses), relative to the target frequency, as functions of the hearing threshold from the first 2-kHz quick auditory filter (qAF) run (top left), threshold-based procedure (top right), second 2-kHz qAF run (bottom left), and 1-kHz qAF run (bottom right). In each panel, the prediction from the regression models (see Figure 7) is plotted as the solid curve with the 95% confidence limits indicated using the dashed curves. The thin dashed horizontal line indicates the summary ERB estimates for listeners with normal hearing (Glasberg & Moore, 1990). Data reported from previous studies are shown as filled symbols.

formulations of the filter's tail. This means that the qAF procedure, as implemented in the current study, may only be sufficient in capturing the auditory filter shape for listeners with hearing threshold less than 50 dB HL at the target frequency.

When comparing auditory filter estimates from various studies (e.g., see Figure 8), one has to be cautious about the differences in the experimental designs across the studies. As pointed out earlier, the formulation of the auditory filter shape is not the same for all studies. Furthermore, these studies also differed in how the stimulus level was determined. Many studies presented stimuli at equal sound pressure levels (e.g., fixed masker spectrum level: Dubno & Dirks, 1989; Glasberg & Moore, 1986; Sommers & Humes, 1993; fixed target level: Baker & Rosen, 2002), whereas the current study presented stimuli at equal sensation levels. This means that higher stimulus levels were used for listener with greater degrees of hearing loss. Because the auditory filter broadens with increasing stimulus levels even for listeners with NH, testing listeners with HI at higher stimulus

levels may introduce a confound between stimulus level and hearing threshold, making it difficult to differentiate the effects of the two factors.

Detection Efficiency and the Effect of Practice

In the current study, listeners' efficiency in detecting a tonal signal in notched noise is described by the parameter $10\log K$, which is associated with the signal-to-noise ratio required after peripheral filtering needed to detect the target tone. It was found that the estimated $10\log K$ was higher for the first 2-kHz qAF run than either the threshold-based procedure (with an average bias of 3.6 dB) or the second 2-kHz qAF run (with an average bias of 1.2 dB). Bland-Altman analyses did not reveal any systematic dependency of the biases on the mean estimates of $10\log K$ (see Figure 4, bottom left). This means that all listeners required slightly higher signal-to-noise ratios to reach detection threshold during the first 2-kHz qAF run while they were still naïve to the experimental

task. This shift of detection efficiency likely reflects practice effects.

It was also observed that the estimated $10\log K$ was lower for the threshold-based procedure than both of the 2-kHz qAF runs, with average biases of 3.6 and 2.4 dB compared to the first and second 2-kHz qAF runs, respectively. These biases may be caused by the procedure difference (i.e., qAF vs. threshold based). Because the notch widths (i.e., g_h and g_l) were varied from trial to trial during the qAF procedure while they were fixed within each experimental block of the threshold-based procedure, the additional trial-by-trial stimulus variability may have caused a slight degradation in detection efficiency. This effect of procedure was distinct from the practice effect because the biases associated with the procedure difference varied across listeners and depended on the value of $10\log K$ (see Figure 6, bottom left). Listeners who were more efficient (lower values of $10\log K$) tended to demonstrate less procedure-related biases. On the other hand, listeners who were less efficient (high values of $10\log K$) exhibited greater biases. For these listeners, $10\log K$ estimates from the threshold-based procedure were lower than those from the qAF procedure.

Significant effects of listener age on $10\log K$ were observed for all three auditory filter estimates at 2 kHz. As age increased, the $10\log K$ increased by about 0.7 dB for every decade of age increase (see the unstandardized regression coefficients in Figure 7, bottom left). This effect of age on detection efficiency has not been reported (e.g., see Sommers & Humes, 1993). Potentially, the large age range and the relatively large number of listeners in the current study supported this novel finding.

Summary

A Bayesian adaptive procedure for the efficient estimation of the auditory filter shapes, the qAF procedure (Shen & Richards, 2013a; Shen et al., 2014), was evaluated using a group of listeners with relatively large ranges of ages and hearing thresholds who were naïve to psychophysical experiments. Robust test–retest reliability was found by comparing the estimated auditory filter parameters across two independent qAF runs tested on separate days (e.g., the auditory filter bandwidths estimated from the two qAF runs were within the tolerance limits). Estimates of the auditory filter using the qAF procedure also agreed closely to the estimates using the traditional threshold-based approach. Moreover, the two qAF runs provided effects of listener age and hearing threshold on the auditory filter parameters (tip sharpness p , tail prominence $10\log w$, and detection efficiency $10\log K$), which agreed well with the threshold-based procedure and with the results of previous studies. Results suggest that the qAF procedure provides a viable means to conduct reliable, time-efficient assessments of auditory frequency selectivity for listeners across a wide age range and with either NH or up to moderate HI.

Acknowledgments

This work was supported by the Hutton Honors College Undergraduate Research Grant at Indiana University awarded to A. B. Kern and the National Institutes of Health Grant R21 DC013406 awarded to co-PIs V. M. Richards and Y. Shen.

References

- American National Standards Institute. (1997). *Methods for calculation of the speech intelligibility index (ANSI S3.5-1997)*. New York, NY: Author.
- Arlinger, S., & Dryselius, H. (1990). Speech recognition in noise, temporal and spectral resolution in normal and impaired hearing. *Acta Oto-Laryngologica. Supplementum*, 469, 30–37.
- Badri, R., Siegel, J. H., & Wright, B. A. (2011). Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms. *The Journal of the Acoustical Society of America*, 129(2), 852–863.
- Baker, R. J., & Rosen, S. (2002). Auditory filter nonlinearity in mild/moderate hearing impairment. *The Journal of the Acoustical Society of America*, 111(3), 1330–1339.
- Baker, R. J., & Rosen, S. (2006). Auditory filter nonlinearity across frequency using simultaneous notched-noise masking. *The Journal of the Acoustical Society of America*, 119(1), 454–462.
- Charaziak, K. K., Souza, P., & Siegel, J. H. (2012). Time-efficient measures of auditory frequency selectivity. *International Journal of Audiology*, 51(4), 317–325.
- Davies-Venn, E., Nelson, P., & Souza, P. (2015). Comparing auditory filter bandwidths, spectral ripple modulation detection, spectral ripple discrimination, and speech recognition: Normal and impaired hearing. *The Journal of the Acoustical Society of America*, 138(1), 492–503.
- Dorr, M., Lesmes, L. A., Elze, T., Wang, H., Lu, Z. L., & Bex, P. J. (2017). Evaluation of the precision of contrast sensitivity function assessment on a tablet device. *Scientific Reports*, 7, 46706.
- Dubno, J. R., & Dirks, D. D. (1989). Auditory filter characteristics and consonant recognition for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 85(4), 1666–1675.
- Dubno, J. R., & Schaefer, A. B. (1991). Frequency selectivity for hearing-impaired and broadband-noise-masked normal listeners. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 43(3), 543–564.
- Dubno, J. R., & Schaefer, A. B. (1992). Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal hearing listeners. *The Journal of the Acoustical Society of America*, 91(4), 2110–2121.
- Fahrmeir, L. (1992). Posterior mode estimation by extended Kalman filtering for multivariate dynamic generalized linear models. *Journal of the American Statistical Association*, 87(418), 501–509.
- Glasberg, B. R., & Moore, B. C. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, 79(4), 1020–1033.
- Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, 47(1–2), 103–138.
- Hopkins, K., & Moore, B. C. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, 130(1), 334–349.

- Jin, S. H., & Nelson, P. B. (2010). Interrupted speech perception: The effects of hearing sensitivity and frequency resolution. *The Journal of the Acoustical Society of America*, 128(2), 881–889.
- Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39(16), 2729–2737.
- Kortlang, S., Mauermann, M., & Ewert, S. D. (2016). Suprathreshold auditory processing deficits in noise: Effects of hearing loss and age. *Hearing Research*, 331, 27–40.
- Laroche, C., Hétu, R., Quoc, H. T., Josserand, B., & Glasberg, B. (1992). Frequency selectivity in workers with noise-induced hearing loss. *Hearing Research*, 64(1), 61–72.
- Leek, M. R., & Summers, V. (1996). Reduced frequency selectivity and the preservation of spectral contrast in noise. *The Journal of the Acoustical Society of America*, 100(3), 1796–1806.
- Leeuw, A. R., & Dreschler, W. A. (1994). Frequency-resolution measurements with notched noises for clinical purposes. *Ear and Hearing*, 15(3), 240–255.
- Lesmes, L. A., Jeon, S. T., Lu, Z. L., & Doshier, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method. *Vision Research*, 46(19), 3160–3176.
- Lesmes, L. A., Lu, Z. L., Baek, J., & Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: The quick CSF method. *Journal of Vision*, 10(3), 17.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2), 467–477.
- Malicka, A. N., Munro, K. J., & Baker, R. J. (2009). Fast method for psychophysical tuning curve measurement in school-age children. *International Journal of Audiology*, 48(8), 546–553.
- Moore, B. C. (1986). Parallels between frequency selectivity measured psychophysically and in cochlear mechanics. *Scandinavian Audiology. Supplementum*, 25, 139–152.
- Moore, B. C. (1995). *Perceptual consequences of cochlear damage*. Oxford, United Kingdom: Oxford University Press.
- Moore, B. C., Vickers, D. A., Plack, C. J., & Oxenham, A. J. (1999). Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *The Journal of the Acoustical Society of America*, 106(5), 2761–2778.
- Myers, J., & Malicka, A. N. (2014). Clinical feasibility of fast psychophysical tuning curves evaluated using normally hearing adults: Success rate, range of tip shift, repeatability, and comparison of methods used for estimation of frequency at the tip. *International Journal of Audiology*, 53(12), 887–894.
- Neher, T., Laugesen, S., Jensen, N. S., & Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *The Journal of the Acoustical Society of America*, 130(3), 1542–1558.
- Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *The Journal of the Acoustical Society of America*, 59(3), 640–654.
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., & Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *The Journal of the Acoustical Society of America*, 72(6), 1788–1803.
- Pepler, A., Munro, K. J., Lewis, K., & Kluk, K. (2014). Repeatability, agreement, and feasibility of using the threshold equalizing noise test and fast psychophysical tuning curves in a clinical setting. *International Journal of Audiology*, 53(10), 745–752.
- Peters, R. W., & Moore, B. C. (1992). Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects. *The Journal of the Acoustical Society of America*, 91(1), 256–266.
- Rosen, S., & Baker, R. J. (1994). Characterising auditory filter nonlinearity. *Hearing Research*, 73(2), 231–243.
- Rosen, S., Baker, R. J., & Darling, A. (1998). Auditory filter nonlinearity at 2 kHz in normal hearing listeners. *The Journal of the Acoustical Society of America*, 103(5, Pt. 1), 2539–2550.
- Sek, A., Alcántara, J., Moore, B. C., Kluk, K., & Wicher, A. (2005). Development of a fast method for determining psychophysical tuning curves. *International Journal of Audiology*, 44(7), 408–420.
- Shen, Y., & Kern, A. B. (2018). An analysis of individual differences in recognizing monosyllabic words under the speech intelligibility index framework. *Trends in Hearing*, 22 2331216518761773.
- Shen, Y., & Richards, V. M. (2013a). Bayesian adaptive estimation of the auditory filter. *The Journal of the Acoustical Society of America*, 134(2), 1134–1145.
- Shen, Y., & Richards, V. M. (2013b). Temporal modulation transfer function for efficient assessment of auditory temporal resolution. *The Journal of the Acoustical Society of America*, 133(2), 1031–1042.
- Shen, Y., Sivakumar, R., & Richards, V. M. (2014). Rapid estimation of high-parameter auditory filter shapes. *The Journal of the Acoustical Society of America*, 136(4), 1857–1868.
- Shen, Y., Zhang, C., & Zhang, Z. (2018). Feasibility of interleaved Bayesian adaptive procedures in estimating the equal-loudness contour. *The Journal of the Acoustical Society of America*, 144(4), 2363–2374.
- Sommers, M. S., & Gehr, S. E. (1998). Auditory suppression and frequency selectivity in older and younger adults. *The Journal of the Acoustical Society of America*, 103(2), 1067–1074.
- Sommers, M. S., & Humes, L. E. (1993). Auditory filter shapes in normal hearing, noise-masked normal, and elderly listeners. *The Journal of the Acoustical Society of America*, 93(5), 2903–2914.
- Song, X. D., Garnett, R., & Barbour, D. L. (2017). Psychometric function estimation by probabilistic classification. *The Journal of the Acoustical Society of America*, 141(4), 2513–2525.
- 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(6), e326–e335.
- Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., & Mott, J. (1985). Speech perception ability and psychophysical tuning curves in hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 77(2), 620–627.
- Stone, M. A., Glasberg, B. R., & Moore, B. C. (1992). Simplified measurement of auditory filter shapes using the notched-noise method. *British Journal of Audiology*, 26(6), 329–334.
- Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *The Journal of the Acoustical Society of America*, 125(5), 3328–3345.
- Summers, V., Makashay, M. J., Theodoroff, S. M., & Leek, M. R. (2013). Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal hearing listeners. *Journal of the American Academy of Audiology*, 24(4), 274–292.
- Vogten, L. L. (1978). Simultaneous pure-tone masking: The dependence of masking asymmetries on intensity. *The Journal of the Acoustical Society of America*, 63(5), 1509–1519.
- Watson, A. B. (2017). QUEST+: A general multidimensional Bayesian adaptive psychometric method. *Journal of Vision*, 17(3), 10.

Appendix (p. 1 of 2)

The Implementation of the Quick Auditory Filter Procedure

The quick auditory filter (qAF) procedure implemented was similar to those described by Shen et al. (2014) with a few exceptions. The MATLAB implementation of the procedure is available at <http://www.indiana.edu/~ahslab/site/software-qaf.html>. A web-based implementation of the qAF procedure (no installation or programming expertise required) is available at <http://www.indiana.edu/~ahslab/qaf>.

Following Rosen et al. (1998), the frequency response of the auditory filter above f_t was given by

$$W_h(g) = (1 + pg)e^{-pg}, \quad (\text{A1})$$

whereas the frequency response of the auditory filter below f_s was given by

$$W_l(g) = (1 - w)(1 + pg)e^{-pg} + w(1 + tg)e^{-tg}. \quad (\text{A2})$$

In Equation A2, the low-frequency skirt of the auditory filter is described by a tip and a tail portion. The parameter p indicates the steepness of the tip portion of the auditory filter, the parameter t indicates the steepness of the tail portion, and the parameter $10\log w$ describes the contribution of the tail portion to the total filter output in decibels. Moreover, the width of the auditory filter was quantified in terms of the equivalent rectangular bandwidth (ERB):

$$\text{ERB} = f_s \left[\int_0^{0.8} W_l(g; p, w) dg + \int_0^{0.8} W_h(g; p) dg \right]. \quad (\text{A3})$$

Distinct from the high-parameter model described by Shen et al. (2014), the steepness of the tip portion was assumed to be identical to that of the low-frequency skirt in the current study (i.e., identical p values for both low- and high-frequency skirts of the filter). Moreover, the steepness of the low-frequency tail was set to $t = 9$. The assumptions on the values of p and t reduced the complexity of the filter shape to containing only two free parameters (the tip sharpness p and the tail prominence $10\log w$). The goal of the simplification was to facilitate the robustness of the qAF procedure. Besides p and $10\log w$, to predict detection thresholds from the auditory filter, an additional parameter $10\log K$ was introduced, which is the signal-to-noise ratio (in decibels) at the output of the auditory filter associated with the detection threshold. Higher $10\log K$ values indicated less efficiency in detecting the tonal target in noise.

Given an auditory filter model, the probability of correct responses was given by

$$\text{PC}(L_M, g_h, g_l; p, w, K) = \gamma + (1 - \gamma) \left\{ 1 + e^{\beta[L_M - \hat{L}_M(g_h, g_l; p, w, K)]} \right\}^{-1}, \quad (\text{A4})$$

where $\gamma = 1/3$ is the chance performance level and $\beta = 1$ is related to the slope of the psychometric function. $\hat{L}_M(g_h, g_l; p, w, K)$ is the masker level at threshold predicted by the auditory filter model:

$$\hat{L}_M(g_h, g_l; p, w, K) = L_T - 10 \log \left[K \int_{g_l}^{g_l+0.25} W_l(g; p, w) dg + K \int_{g_h}^{g_h+0.25} W_h(g; p) dg \right], \quad (\text{A5})$$

where L_T is the target level expressed in dB SPL.

In the qAF procedure, the estimation of the parameters p , $10\log w$, and $10\log K$ was achieved by updating a posterior parameter distribution following each behavioral response. The posterior parameter distribution was modeled as a multivariate normal distribution in the three-dimensional parameter space, defined by the axes along p , $10\log w$, and $10\log K$. Following the i th qAF trial, the posterior parameter distribution was represented by its mean ϕ_i and covariance matrix $\mathbf{P}_i$. Depending on whether the response collected on the i th trial was correct ($r = 1$) or incorrect ($r = 0$), ϕ_i and $\mathbf{P}_i$ were approximated using extended Kalman filtering (e.g., Fahrmeir, 1992):

$$\phi_i = \phi_{i-1} + \mathbf{K}(r - \mu) \quad (\text{A6})$$

Appendix (p. 2 of 2)

The Implementation of the Quick Auditory Filter Procedure

and

$$\mathbf{P}_i = \tilde{\mathbf{P}}_{i-1} - \mathbf{K} \cdot \mathbf{J} \cdot \tilde{\mathbf{P}}_{i-1} \quad (\text{A7})$$

where

$$\mathbf{K} = \tilde{\mathbf{P}}_{i-1} \cdot \mathbf{J}' \cdot [\mathbf{J} \cdot \tilde{\mathbf{P}}_{i-1} \cdot \mathbf{J}' + \mu(1-\mu)]^{-1}, \quad (\text{A8})$$

$$\mathbf{J} = \left\{ \frac{\partial \text{PC}}{\partial p} \quad \frac{\partial \text{PC}}{\partial 10 \log w} \quad \frac{\partial \text{PC}}{\partial 10 \log K} \right\}, \quad (\text{A9})$$

and

$$\mu = \text{PC}(L_M, g_h, g_i; \varphi_{i-1}). \quad (\text{A10})$$

The covariance matrix $\tilde{\mathbf{P}}_{i-1}$ is the same as $\mathbf{P}_{i-1}$, except that its diagonal elements, that is, variances for p , $10 \log w$, and $10 \log K$, are increased by 5% (i.e., the diffusion factor). Moreover, the qAF procedure limited the estimated p values to be between 20 and 70 (dimensionless), the estimated $10 \log w$ values to be between -60 and 0 dB, and the estimated $10 \log K$ values to be between -10 and 10 dB, which defined the boundaries of the parameter space. The Kalman filtering procedure (Equations A6 and A7) was only carried out when φ_i was within the boundaries. The use of the diffusion factor and the limits on the parameter values were implemented to improve the stability of the qAF procedure (see Shen et al., 2014, for detailed discussions).

Before each qAF run, the prior parameter distribution was defined as $\varphi_0 = \{40, -30, 5\}^T$ and $\mathbf{P}_0 = \text{diag}\{1600, 1600, 400\}$ (prior Configuration 1) for the first 18 listeners, $\varphi_0 = \{40, -15, 5\}^T$ and $\mathbf{P}_0 = \text{diag}\{1600, 1600, 4\}$ (prior Configuration 2) for the next 15 listeners, and $\varphi_0 = \{40, -15, 5\}^T$ and $\mathbf{P}_0 = \text{diag}\{1600, 1600, 25\}$ (prior Configuration 3) for the last 17 listeners. The prior distribution was adjusted twice during data collection, because improved a priori knowledge about the parameter distribution improved as more data were collected. Between-subjects analyses of variance showed no significant effect of different prior configurations on the parameter estimates for the two 2-kHz qAF runs: first 2-kHz qAF run: $F(2, 47) = 1.89$, $p = .163$ for tip sharpness p ; $F(2, 47) = 1.15$, $p = .327$ for tail prominence $10 \log w$; $F(2, 47) = 1.18$, $p = .316$ for detection efficiency $10 \log K$; second 2-kHz qAF run: $F(2, 47) = 1.05$, $p = .359$ for tip sharpness p ; $F(2, 47) = 1.83$, $p = .172$ for tail prominence $10 \log w$; $F(2, 47) = 0.25$, $p = .780$ for detection efficiency $10 \log K$. Therefore, the parameter estimates from the qAF procedure were not very sensitive to variations in prior configurations. During data analysis, data were collapsed across prior configuration groups.

The stimulus presented on each qAF trial was identified as the combination of L_M , g_h , and g_i that would lead to the greatest expected reduction in entropy for the posterior parameter distribution (e.g., Kontsevich & Tyler, 1999; Lesmes et al., 2006, 2010; Shen & Kern, 2018; Shen & Richards, 2013a, 2013b; Shen et al., 2014; Watson, 2017).