

The Human Cochlear Mechanical Nonlinearity Inferred via Psychometric Functions

Lance Nizami

Correspondence: nizami2@att.net
Independent Research Scholar, Palo Alto, CA, USA

Abstract

Background: Schairer and colleagues hypothesized that the slope of any psychometric function for forward-masked probe-tone detection depends upon the standard deviation of an external Gaussian input distribution of probe-tone intensities, which in turn reflects the coupling of an internal Gaussian output distribution to the cochlear mechanical nonlinearity. The latter was postulated to have two conjoined branches, straight lines of differing slopes. Hence, just two possible standard deviations were predicted for the external input distributions, i.e., one for each branch of the nonlinearity – and therefore two different slopes of psychometric functions for forward-masked probe-tone detection. To confirm the latter, Schairer and colleagues obtained psychometric functions for the detection of forward-masked probe-tones.

Methods: Such psychometric functions were already available, for detection thresholds of unprecedented precision which had been found as a function of either (1) time gap between probe-tone and same-frequency constant-intensity forward-masker (“recovery”), or (2) same-frequency forward-masker intensity at fixed masker-probe time gap (“growth”). Those psychometric functions were re-analyzed here because the model of Schairer and colleagues can be extended such that the hypothesized relation of a probe-tone’s psychometric-function slope to its detection threshold can be specified as an equation, through which psychometric-function slope becomes proportional to the cochlear nonlinearity’s own rate-of-change with intensity.

Results: The cochlear nonlinearity’s rate-of-change, for “recovery” data, follows an angle-shape, whereas for “growth” data it declines as a single power function; it does so also for re-analyzed probe-tone-detection thresholds of Schairer and colleagues. The equations for rates-of-change were integrated to give the cochlear nonlinearities themselves, each characterized by a single unknown parameter. The parameter’s possible values were implied by comparing the nonlinearity’s inferred rates-of-change in man to those measured in animals. Altogether, then, the human cochlear nonlinearities inferred from “recovery” have a distinct but smooth bend between two branches, a steep low-intensity branch and a shallow high-intensity branch, whereas those inferred from “growth” resemble the smoothly decelerating nonlinearities observed for animals.

Conclusions: Extension of the model of Schairer and colleagues results in credible cochlear nonlinearities in man, suggesting that forward-masking provides a non-invasive way to infer the human mechanical cochlear nonlinearity.

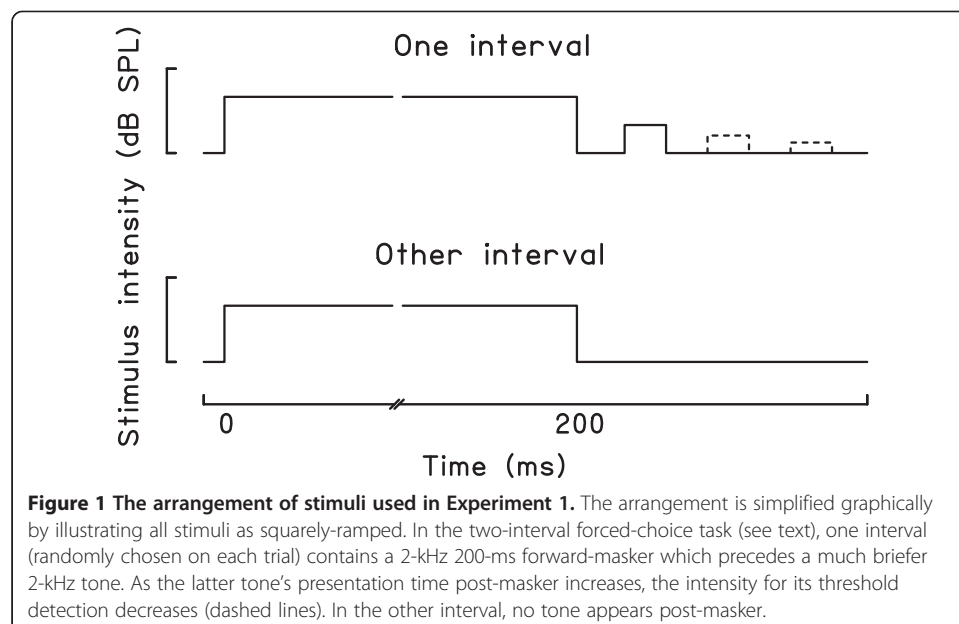
Keywords: Cochlear mechanical nonlinearity; Psychometric function; Detection threshold; Forward-masking

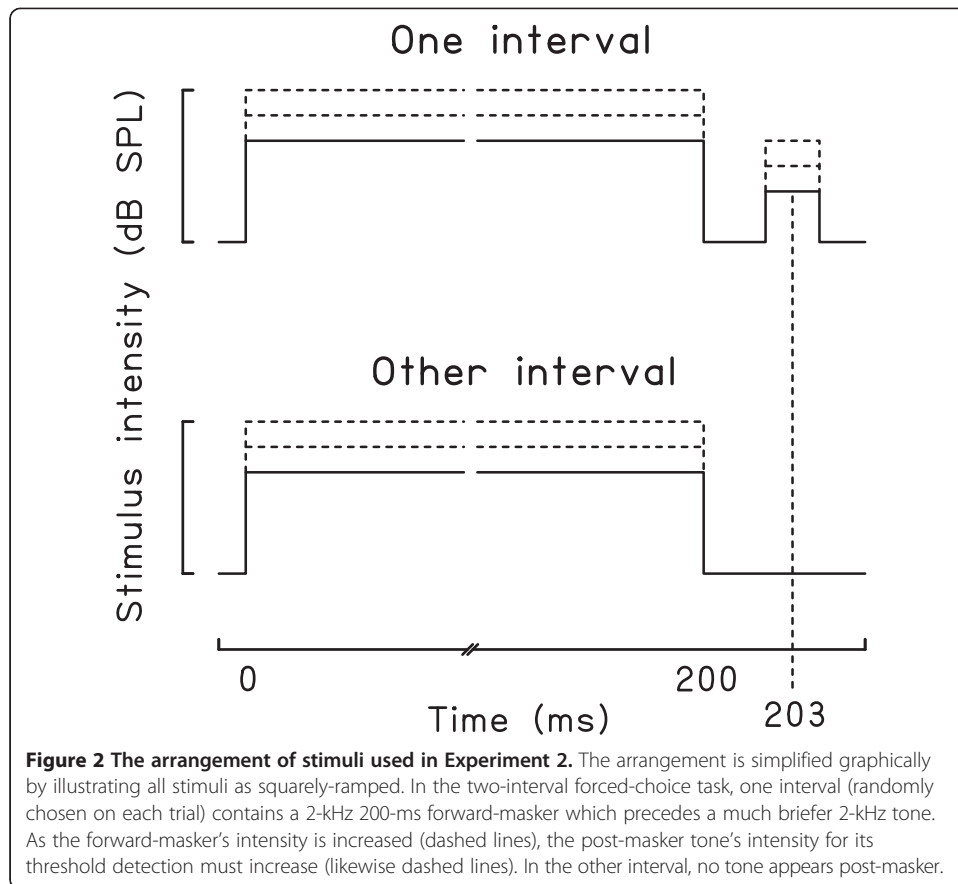
Background

Forward-masking and the psychometric function

When an auditory “probe” stimulus is preceded by a “forward-masker” stimulus, which is usually longer, the probe’s detection threshold is typically raised above that found in quiet. This phenomenon is called “forward masking”, in contrast to simultaneous masking, in which masker and probe overlap in time. When forward-masker intensity, forward-masker and probe duration, and interval between forward-masker and probe are constant, the greatest elevation of the probe’s threshold under forward-masking occurs when the center of the frequency spectrum of the forward-masker coincides with the center of the frequency spectrum of the probe. The simplest example is that of a forward-masker and probe-tone of the same single sine-wave frequency. For a given probe-tone frequency and duration, it is well-established that the probe-tone’s detection threshold generally decreases with increasing time between the end of the forward-masker and the start of the probe-tone, and generally increases with increasing forward-masker intensity. Figure 1 shows the former case, which is used in Experiment 1 here; Figure 2 shows the latter case, which is used in Experiment 2 here.

The experimental method used in such listening tasks is typically two-interval, two-alternative forced-choice. On each “forced choice” trial, two successive intervals contain the forward-masker but only one (randomly chosen) contains the probe-tone. Listeners are required to specify which interval contains the probe-tone. For particular probe-tone and forward-masker frequencies and durations, and for a particular forward-masker intensity and a particular masker-to-probe-tone time gap (hereafter called “masker-probe time gap”), a listener’s incidence of correctly specifying the interval containing the probe-tone (described by a percentage-correct score, or divided by 100 as a probability) can be plotted as a function of probe-tone intensity (in dB SPL),





yielding a set of points. A curve fitted to those points is dubbed the psychometric function for probe-tone detection. That psychometric function can be approximated algebraically as an ogive (S-curve); one way to do that is through the well-established curve-fitting procedure called Probit Analysis ([1]; more details below). The probe-tone's intensity at the ogive's midpoint is typically taken as the probe-tone's putative detection threshold. Many examples of psychometric functions will be presented here.

The cochlear compressive nonlinearity

Within the last decade, forward-masking has been used to examine the possible operation, in humans, of something examined physiologically in animals and which, over three decades, has become the dominant concept in auditory mechanics and in hearing psychology: the cochlear compressive mechanical nonlinearity. Consider the following brief summary based upon reviews by Ulfendahl [2] and Robles and Ruggero [3]. Understanding the cochlear nonlinearity requires first understanding the concept of a nonlinear system. In a nonlinear system, output does not vary in direct proportion to input. The cochlear nonlinearity itself is the stimulus-driven motion of the organ of Corti, a surface of cells which curls within the snail-shaped peripheral electromechanical hearing organ, the cochlea, and which is sometimes referred to by its internal support structure, the basilar membrane. The basilar membrane vibrates (with some time

delay, of course) in response to the pressure oscillations of an auditory stimulus (a “sound wave”). The basilar membrane’s *peak* vibration amplitude can be expressed in decibels by taking its logarithm with respect to a reference level (similar to the expression of the sound-wave’s intensity as root-mean-square vibration amplitude in decibels sound-pressure-level (dB SPL)).

Linearity of basilar-membrane vibration has been used colloquially to mean that a change of 1 dB in stimulus intensity causes a change of 1 dB in basilar-membrane vibration. That is, basilar-membrane output is a power function of exponent 1 of stimulus input. Consider a pure tone as the stimulus. Empirically, the basilar membrane’s greatest response in dB to a pure tone, defined as the response at the “characteristic frequency” (CF) place of the pure tone on the basilar membrane, does *not* increase in direct proportion to the tone’s intensity in dB SPL, for the larger part of the ear’s useful listening range. Rather, for *moderate* tone intensities in the basal (i.e., high-CF) cochlear turn of chinchillas, Guinea pigs, and cats, a ratio of output to input of < 1 dB/dB, called “the compressive nonlinearity”, is seen, with some output/input ratios being as low as 0.2 dB/dB. (Experiments are ongoing for the apical [i.e., low-frequency] region, which is harder to study.) An illustration of cochlear nonlinearity will soon be presented.

At *low* tone intensities, however, the input-output response is linear. Such responses, of 1 dB/dB, are also suggested for *sufficiently higher-than-moderate* intensities of pure tones; both theoretically and empirically, the input–output relation steepens above 80 dB SPL, its slope approaching unity at 90–100 dB SPL. Nonetheless, some studies indicate highly compressive growth right up to 100 dB SPL (Guinea pig, chinchilla). The high-intensity regime is not relevant to the model considered here, and will not be mentioned further. Similarly, the compression zone’s *lower* limit is unclear; for example, in the chinchilla it can be roughly 40 dB SPL, but in the Guinea pig it can be roughly 50 dB SPL. Whether the basilar membrane’s input-output function varies across species is an ongoing debate. It is agreed, however, that nonlinearity is only present at or near the basilar-membrane locus of the CF. When the frequency of the stimulating tone is lower than about 0.7 of the CF, the vibrations there grow linearly, at least for basilar-membrane loci at the cochlear base. The present paper concerns only basilar-membrane loci at the respective CF’s.

Relating the cochlear nonlinearity to the psychometric function for forward-masked probe-tone detection: Schairer et al. [4]

Schairer et al. [4] related the cochlear nonlinearity to the psychometric function for forward-masked probe-tone detection. Evidence supporting the Schairer et al. [4] model was later pursued by Schairer et al. [5].

The model of Schairer et al. [4]

Schairer et al. [4] credit various aspects of their model to Plack and Oxenham [6], specifically, the chosen form of the cochlear nonlinearity, and the use of multiplicative internal noise, or as Plack and Oxenham [6, p. 1599] put it, “assuming that [probe-tone detection] threshold corresponds to a constant internal signal-to-masker ratio”. Here, “internal masker” of course refers to the internal response to the forward-masker.

Schairer et al. [4, p. 1561] postulated that estimation of forward-masked probe-tone detection threshold involves an “underlying distribution of input signal levels” (i.e., an external, input distribution of probe-tone *intensities*). The latter is assumed to be Gaussian, and to reflect an internal Gaussian “output distribution”, one of fixed standard deviation regardless of its mean value. That is, the input distribution depends upon the shape and standard deviation of the required output distribution. In the absence of knowledge of the latter standard deviation, as well as of the form of the cochlear nonlinearity in man, in the input distribution can be established only over the course of a forward-masked threshold-detection experiment. The input distribution’s standard deviation reflects the slope of the psychometric function for probe-tone detection. (The mean value of the distribution of probe-tone intensities is one candidate for the probe-tone detection threshold; as used presently, see below). Schairer et al. [4] said nothing about how forward-masking determines the mean value of either the input or the output distribution (given that one mean value determines the other).

The Schairer et al. [4] model involved a particular algebraic simplification for the nonlinearity. That is, for tone intensities of 0-35 dB SPL, Schairer et al. [4] hypothesized that basilar-membrane vibration amplitude in decibels follows stimulus vibration in decibels SPL. But for tones of higher SPLs, an increase of 1 dB SPL was assumed to produce an increase of “around 0.2” decibels of basilar-membrane vibration ([4, p. 1560], after [7]). The cochlear nonlinearity thus hypothetically consisted of two line segments joined at a single point. Figure 3 shows the Schairer et al. [4] model. The choice of 0.2 and the abruptness of the transition between “linear” and “compressive” stages in man are actually both contentious; these issues are important enough to receive their own section in the Discussion.

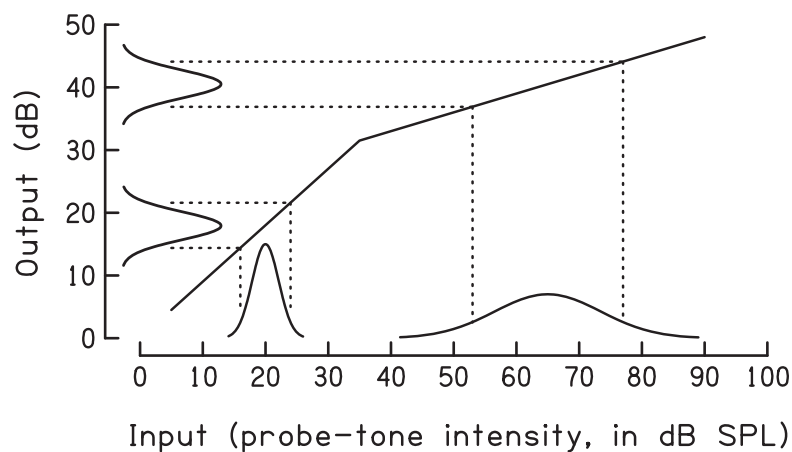


Figure 3 The Schairer et al. [4] model. The model shows the effect of the cochlear mechanical nonlinearity upon the distributions of probe-tone intensities involved in probe-tone detection, and hence upon the slopes of the psychometric functions for forward-masked probe-tone detection (see text). Note well that the true amplitudes of the probability density functions are not dB or dB SPL, as might appear from the graph, but rather are probability density, imagined as the label of a z-axis rising perpendicularly out of the page from {0,0}. As such, the shown probability density functions are projections upon the input/output plane of the graph. Also, for illustration's sake, the input distributions (and hence the output distributions) shown here are at least twice as wide as will be eventually implied from the present empirical psychometric functions.

Testable predictions of the model of Schairer et al. [4]

The distribution of input probe-tone intensities in decibels SPL is hypothetically (1) Gaussian with fixed standard deviation for the “linear” section of the Schairer et al. [4] putative cochlear nonlinearity (Figure 3, for probe-tone intensities < 35 dB SPL), (2) Gaussian with a larger fixed standard deviation for the “compressive” section of the Schairer et al. [4] putative cochlear nonlinearity (Figure 3, for probe-tone intensities > 35 dB SPL). The consequent hypothetical psychometric functions would be steep for low probe-tone intensities (such as 5-35 dB SPL), but would be shallower for moderate probe-tone intensities (i.e., 35-90 dB SPL). In short, within the Schairer et al. [4] model, the slope of the psychometric function for probe-tone detection (measured, for example, at that function’s midpoint) should have either of just two possible values. Unfortunately, those slopes are unknown, because a crucial number, the standard deviation of the hypothesized Gaussian output distribution, is unknown. Even the intensity at which those two possible psychometric-function slopes change, one to another, is unknown. Schairer et al. [4] set that bend in the nonlinearity at 35 dB SPL (Figure 3), as Plack and Oxenham did [6, p. 1604]. Nonetheless, Schairer et al. [4] did not assume the bend-point to be the same for every ear.

The Schairer et al. [4] model was tested by Schairer et al. [4] and later by Schairer et al. [5]. The results of those experiments provide a vital context for the present experiments.

Further background: necessary details of the experiments conducted by Schairer et al. [4] and later by Schairer et al. [5]

A forward-masker can be used to change the detection threshold of a pure-tone probe stimulus without changing the latter’s duration or its frequency; after all, such changes would confuse things by changing the shape of the nonlinearity at a given CF [2,3]. As a test of their model, then, Schairer et al. [4] manipulated the detectability of a probe-tone of fixed frequency and duration by using forward-maskers of different intensities presented at a fixed masker-probe time gap. In particular, Schairer et al. [4] identified two experimental conditions which hypothetically offered complementary psychometric functions for probe-tone detection, as follows. Psychometric functions were obtained in the “variable-signal” condition where “signal” meant “probe-tone”, at each of a broad range of forward-masker intensities. For each of the latter, Schairer et al. [4] plotted the experimental listener’s percentage-correct scores versus probe-tone intensity, x , in dB SPL, and then fitted to those points a smooth psychometric function represented by an equation (although their method of fit differs from the present one; see the Discussion). Alternatively, in the “variable-masker” condition, where “masker” refers to the forward-masker, the experimenters chose a number of fixed *probe-tone* intensities covering a broad range of SPLs, and then constructed a psychometric function for the detection of each probe-tone, by varying the *forward-masker* intensity. (Note that this is not the design of the present Experiment 2). By increasing the forward-masker intensity, the percentage-correct score for identification of a fixed-intensity probe-tone would decrease from near-certainty to near-chance, giving a *reversed* ogive when percentage-correct is plotted as a function of forward-masker intensity. Later, Schairer et al. [5] tested hearing-impaired listeners as well as normal-hearing listeners. Hearing-impaired

listeners were presumed to show linearized responses to pure tones, which is certainly the case in animals postmortem or with acoustic overstimulation [2, pp. 350-353] [3, pp. 1334-1335]. The psychometric functions of hearing-impaired listeners should therefore remain steep regardless of forward-masker intensity, rather than becoming shallower for a fixed masker-probe time gap and increased forward-masker (and hence probe-tone) intensity.

The variable-signal experiments, when conducted using normal-hearing subjects, generally supported the model, whereas the variable-signal experiments in hearing-impaired subjects, and the variable-masker experiments generally, were somewhat equivocal. Strong individual differences in performance were apparent across subjects, especially for the hearing-impaired listeners and for the variable-masker experiments.

Altogether, the Schairer et al. [4] model can hardly be dismissed. Presently, data interpretation is simplified by not using hearing-impaired subjects as in Schairer et al. [5] or variable-masker conditions as in Schairer et al. [4] and Schairer et al. [5]; such situations were either impractical or needless (see below).

A necessary methodological prelude: overview of the present experiments

The Schairer et al. [4] and Schairer et al. [5] results suggested that their model could be correct, but were not conclusive. The Schairer et al. [4] model can be tested further, using data that reliably document the slopes of the psychometric functions for probe-tone detection thresholds over a broad range of probe-tone detection thresholds. Empirically, for the small furred mammals whose cochlear nonlinearities appear in the literature, the degree of cochlear compression monotonically increases with increasing stimulus intensity for low-to-moderate stimulus intensities [2,3], changing smoothly from not compressive, or lightly compressive, to very compressive, rather than having the bisegmented shape of Figure 3. That is, the cochlear nonlinearity gets flatter and, as such, within the Schairer et al. [4] model the widths of the psychometric functions for probe-tone detection will monotonically increase, such that the slopes of those psychometric functions will monotonically decrease^a.

Now to the present Experiments 1 and 2. The first of the experiments was thoroughly described in a paper [8] which was, nonetheless, prematurely brief. The paper preceded the finishing of a dissertation [9] which contained the details of Experiment 1, but not of Experiment 2. The latter had also preceded the finishing of [9] but was not mentioned therein, as it was purely auxiliary and, in fact, it was not described until a recent Proceeding [10]. Ironically, neither of the two experiments was originally intended to test any model of the effects of cochlear nonlinearity; on the contrary, one experiment was a detailed study of the course of recovery from forward-masking, and the other explored the increase in forward-masking with increase in forward-masker intensity.

Overview of experiments 1 and 2

In describing the present experiments, it is necessary to return momentarily to Schairer et al. [4]. Their variable-signal condition *resembles* the conditions used presently, in that the probe-tone intensity (rather than the forward-masker intensity) is what is varied to provide percentages-correct to which to fit a psychometric function. Each such fitted psychometric function has a slope at any point along itself, a slope which can be

described in units of cumulative probability (that is, percentage-correct divided by 100) per decibel, labeled “1/dB” from here on. Schairer et al. [4] hypothesized that the slope of the psychometric function of a forward-masked probe-tone depends upon the probe-tone detection threshold. To test their hypothesis, they had elevated the probe-tone’s threshold to different degrees by fixing the masker-probe time gap at relatively small values – gaps on the order of their short probe-tone’s duration – and then using different intensities of the forward-masker. Detection threshold was then plotted versus forward-masker intensity in dB SPL, producing a plot of growth of forward-masking.

However, there is another approach to obtaining psychometric functions for detection of probe-tones under forward-masking. That is, it is well-known that under a relatively long and intense forward-masker, such as a 200-ms (milliseconds) tone of 90 dB SPL or higher, the detection threshold of a shorter tone of the same frequency can be elevated by many tens of decibels (citations in [8]). Thus elevated, the probe-tone’s detection threshold may not “recover” to its value in quiet for several hundred milliseconds, perhaps even several seconds. As masker-probe time gap increases, probe-tone detection threshold generally declines; this plot of recovery appears to be generally monotonic. This would seem advantageous experiments-wise, because as the probe-tone intensity changes from moderate to low, the cochlear nonlinearity itself also changes in a generally monotonic fashion, transiting from shallower to steeper [2,3]. Hence, if the slope of the psychometric function for probe-tone detection depends upon the cochlear nonlinearity, as hypothesized by Schairer et al. [4], then the psychometric functions for probe-tone detection should generally steepen with those probe-tones’ presentation times postmasker. This was Experiment 1.

Experiment 2 obtained probe-tone detection thresholds under what Schairer et al. [4] called the “variable-signal” condition, using a fixed masker-probe time gap, as in Schairer et al. [4] or Schairer et al. [5], but for more forward-masker intensities than used in [4] or in [5].

Methodology common to the present experiments

Compared to Schairer et al. [4] and Schairer et al. [5], fewer subjects, run singly, were involved in much longer experiments, the purpose being to produce a few plots of recovery from forward-masking, or of growth of forward-masking, but in greater detail. The emphasis was also to describe the probe-tone-detection psychometric functions themselves using far greater numbers of percentages-correct (see below) than ever done before. Psychometric functions have long been valued as indicators of the detection process, but the literature contains surprisingly few psychometric functions for the detection of forward-masked probe-tones. Further, the precision of those examples has been low, due to the relatively low numbers of forced-choice trials employed. The issue of precision is crucial, and is covered in a separate Discussion section.

In the first experiment, three experimental subjects were run, and in the second, just two; time, finances, and space altogether allowed no more. A double-walled soundproof chamber was employed in Experiments 1 and 2. It only allowed space for one subject at a time, but sitting comfortably and without claustrophobia. The subjects of Experiment 1 were described in [8]. In Experiment 2, Subject 1A was in fact Subject 1 of Experiment 1. Subject 2A was a recent university graduate (male, age 24) who was paid hourly for

participation. All subjects were given extensive practice before formal data collection ensued, and became quite adept.

All stimuli were played to the right ear. Each probe-tone detection threshold was found using blocks of 100 self-paced two-interval two-alternative forced choices. In both Experiment 1 and Experiment 2, the probe-tone intensity and the forward-masker intensity (and the masker-probe time gap) were all fixed during any block of forced-choices. This is the conservative and well-tried “method of constant stimuli”. To obtain different percentages-correct for fitting to a psychometric function, probe-tone intensity was changed across blocks, but adjacent intensities never differed by more than 2 dB.

The number of different probe-tone intensities used to determine a probe-tone detection threshold increased with the evident width of the psychometric function, as assessed block-by-block by the experimenter from the subject’s percentages-correct, covering the psychometric function evenly from roughly 55% correct to 95% correct and hence assuring a relatively constant error in the inferred detection thresholds. To construct a psychometric function, the experimental subject’s percentage-correct scores are divided by 100, and the results are fitted to an S-shaped “ogive”. The resultant psychometric functions for probe-tone detection were indeed typical ogives, like those found for “variable-signal” conditions in Schairer et al. [4] and later in Schairer et al. [5], but not the reversed-S curves found by Schairer et al. [4] and later by Schairer et al. [5] for “variable-masker” conditions. When a sufficient range of percentages-correct had been obtained to confidently fit using a psychometric function, either the masker-probe time gap (Experiment 1) or the forward-masker intensity (Experiment 2) was changed.

The method of constant stimuli is ponderous, and strongly dissuades the use of hearing-impaired listeners, who require closer attention, and more time, than normals. They may also frustrate more easily. Indeed, the present experiments were intended only for normative responses in unimpaired listeners. In the first listening block for any new masker-probe time gap t (Experiment 1) or any new forward-masker intensity (Experiment 2), the probe-tone’s intensity was set so high that the listener made few mistakes. The probe-tone’s intensity was then slowly lowered over successive blocks.

Probit analysis

For each employed masker-probe time gap t (Experiment 1) or forward-masker intensity (Experiment 2), the psychometric function, the probe-tone detection threshold, and the psychometric-function slope were all obtained through Probit Analysis [1]. The employed ogive is the cumulative integral of a Gaussian probability density function of mean value μ (in dB SPL) and standard deviation σ (in dB). For two-interval two-alternative forced choice, the ogive is

$$P(x) = \frac{1}{2} + \frac{1}{\sigma\sqrt{8\pi}} \int_{-\infty}^x \exp \frac{-(x-\mu)^2}{2\sigma^2} dx, \quad -\infty \leq x \leq \infty, \quad 0.5 \leq P(x) \leq 1 \quad (1)$$

(see [8]). $P(\mu) = 0.75$, such that μ yields the midpoint of the ogive, and is taken as the stimulus-detection threshold. At μ , the slope of the ogive has its greatest value which,

due to symmetry, is its one *unique* value. That slope, in units of cumulative probability per decibel, is

$$\text{slope} = \left. \frac{dP(x)}{dx} \right|_{x=\mu} = \frac{1}{\sigma\sqrt{8\pi}} \text{ 1/dB} \quad (2)$$

Probit Analysis allows confidence intervals to be computed for each threshold [1]. The 95% (or better) confidence intervals are convenient error bars. When those bars do not overlap for neighboring probe-tone detection thresholds, the respective thresholds can be considered to differ significantly. The present probe-tone detection thresholds show confidence intervals which are smaller by far than any error bars shown in the literature, either for probe-detection thresholds of individuals or of groups. Altogether, the obtained psychometric functions offer an unprecedented degree of precision.^b This is no idle boast, but its justification is detailed and hence relegated to the Discussion.

The nature and advantage of the particular probe-tone stimulus

In both experiments here, the forward-masker and the probe-tone were both of 2 kHz. The forward-masker was 200 ms long (not including end-ramps). The probe-tone's duration was determined by its shape. It was shaped by a Gaussian envelope, which provides the smallest theoretical spread in the frequency actually experienced at the basilar membrane [11,12]. (The same Gaussian envelope was applied to the start and to the end of the forward-masker). The Gaussian had a standard deviation of 0.5 ms, equal to the tone's sine-wave period. The amplitude was set to zero at ± 3 standard deviations, giving an actual duration of $6 \times 0.5 = 3$ ms. The probe-tone's energy is relatively narrowly centered on the basilar membrane; the relative spectral energy density has a single lobe spanning 1.517-2.483 kHz at 10 dB below its maximum (noted in [8]; illustrated in [13]). The probe did not subjectively resemble a click or a narrowband noise; rather, it sounded like a bubbling or clapping, partway between click and tone.

Gaussian-shaped probe-tones have a further special advantage, in that they allow the time gap t between the probe-tone and the termination of the forward-masker to be meaningfully specified, i.e., as the interval between the beginning of the forward-masker's terminal decline (ramping-down) and the peak of the probe-tone's envelope. For non-Gaussian-shaped tones typical of the literature, the masker-probe time gap is usually measured from the termination of the forward-masker to the *start* of the probe-tone. But the interval from forward-masker termination to the *end* of the probe-tone could be used just as well when forward-masker and probe-tone do not overlap, as could the interval to the *middle* of the probe-tone, and so on. The use of Gaussian-shaped probe-tones reduces such ambiguity. However, the forward-masker and the probe-tone do physically overlap at $t = 0$, and for 2 ms or so thereafter [9, Figure 36]. t was always an integer multiple of the 0.5 ms period.

Method and results of experiment 1: threshold recovery from forward-masking

The method has been described in great detail elsewhere [8]. Briefly, the forward-masker's intensity was set at 97 dB SPL, which would produce substantial elevation of the probe-tone's detection threshold. The psychometric function for probe-tone detection was obtained at various post-masker time gaps, t , which followed a randomized search pattern intended to cover $0 \leq t \leq 40$ ms in a manner that was customized for each subject, in view

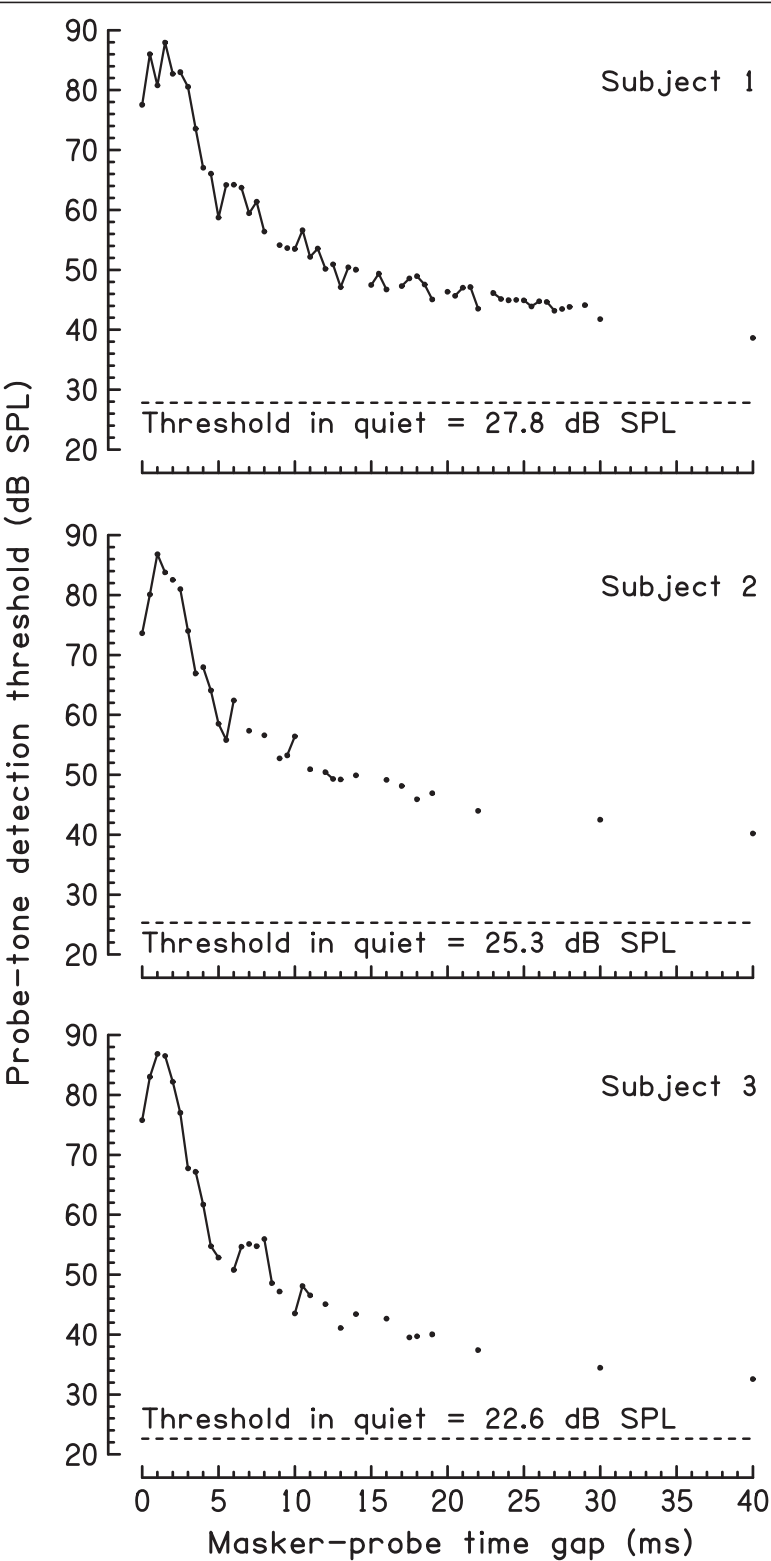


Figure 4 Experiment 1: the detection thresholds for the forward-masked 2-kHz Gaussian-shaped probe-tone.

of individual differences that quickly became apparent. Thus, subjects did not all have the same t 's; they were chosen to be closer together as $t \rightarrow 0$, so that t 's were $\frac{1}{2}$ ms apart for all subjects for $t \leq 6$ ms, but could be 1 ms apart, or even greater integer multiples of $\frac{1}{2}$ ms, for $t > 6$ ms. Subject 1 had more time available than the other subjects, and consequently produced detection thresholds at more time gaps than did Subjects 2 and 3.

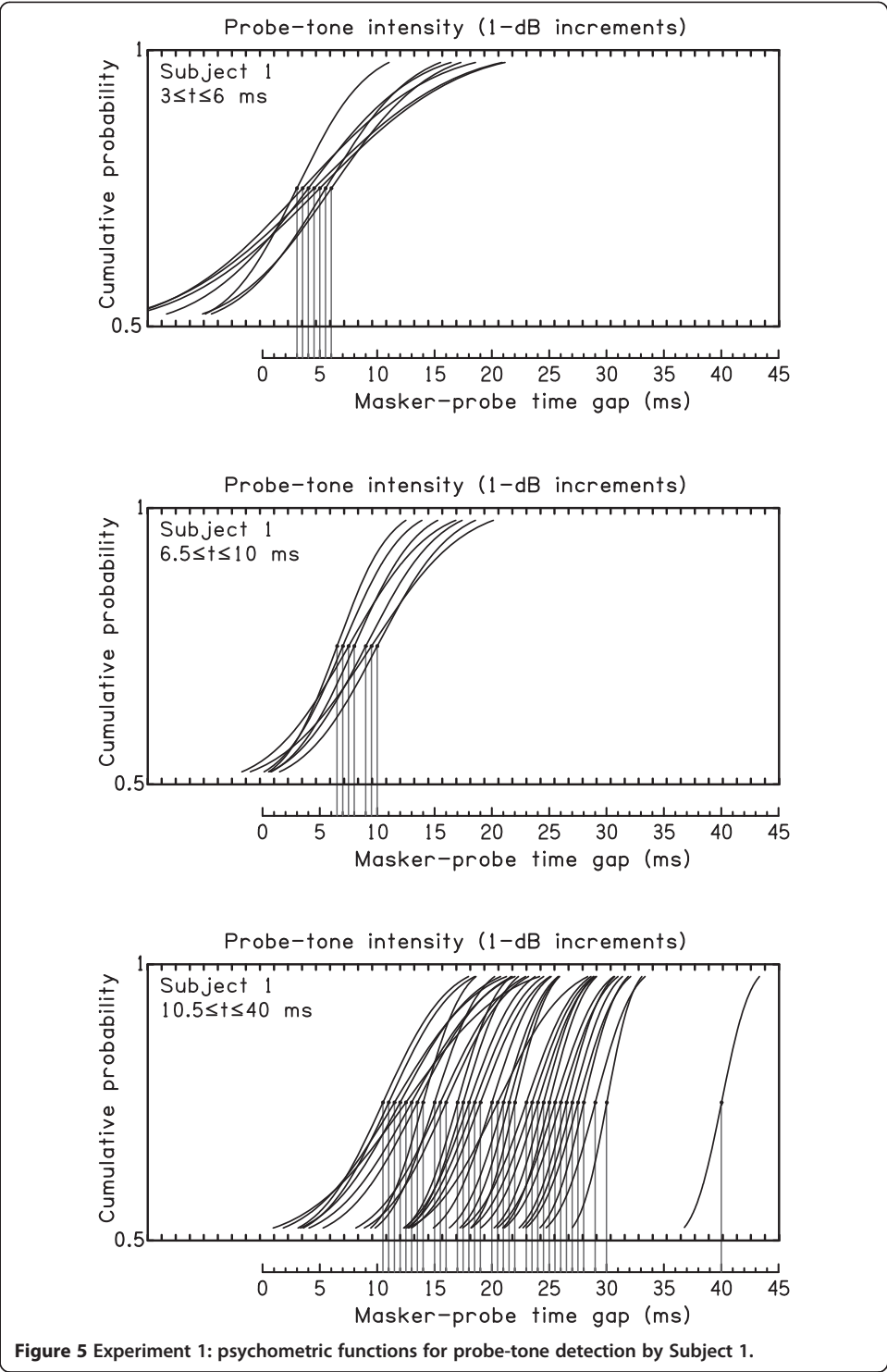
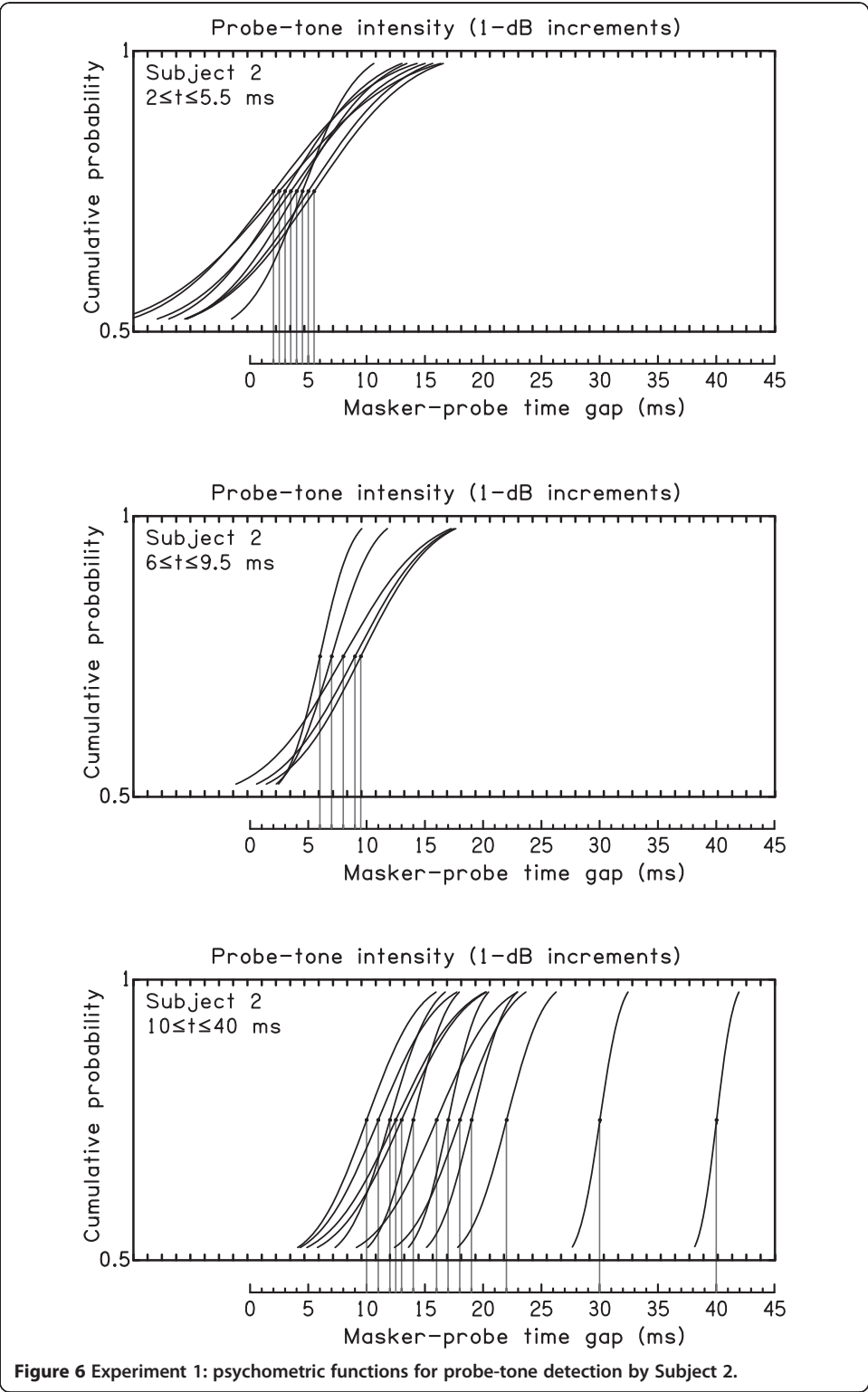


Figure 4 shows the actual probe-tone detection thresholds for Subjects 1–3. The error bars were removed for the sake of reducing clutter; the interested reader can find them elsewhere ([8, Figure 1]; they are very small indeed, typically 1 dB or less). Such fine resolution cannot be found elsewhere. Points that were ½ ms apart in time gap



and that also differed significantly in detection threshold are joined by lines in Figure 4. These differences are not apparent in other publications (see Discussion), and reflect the greater precision of the present detection thresholds.

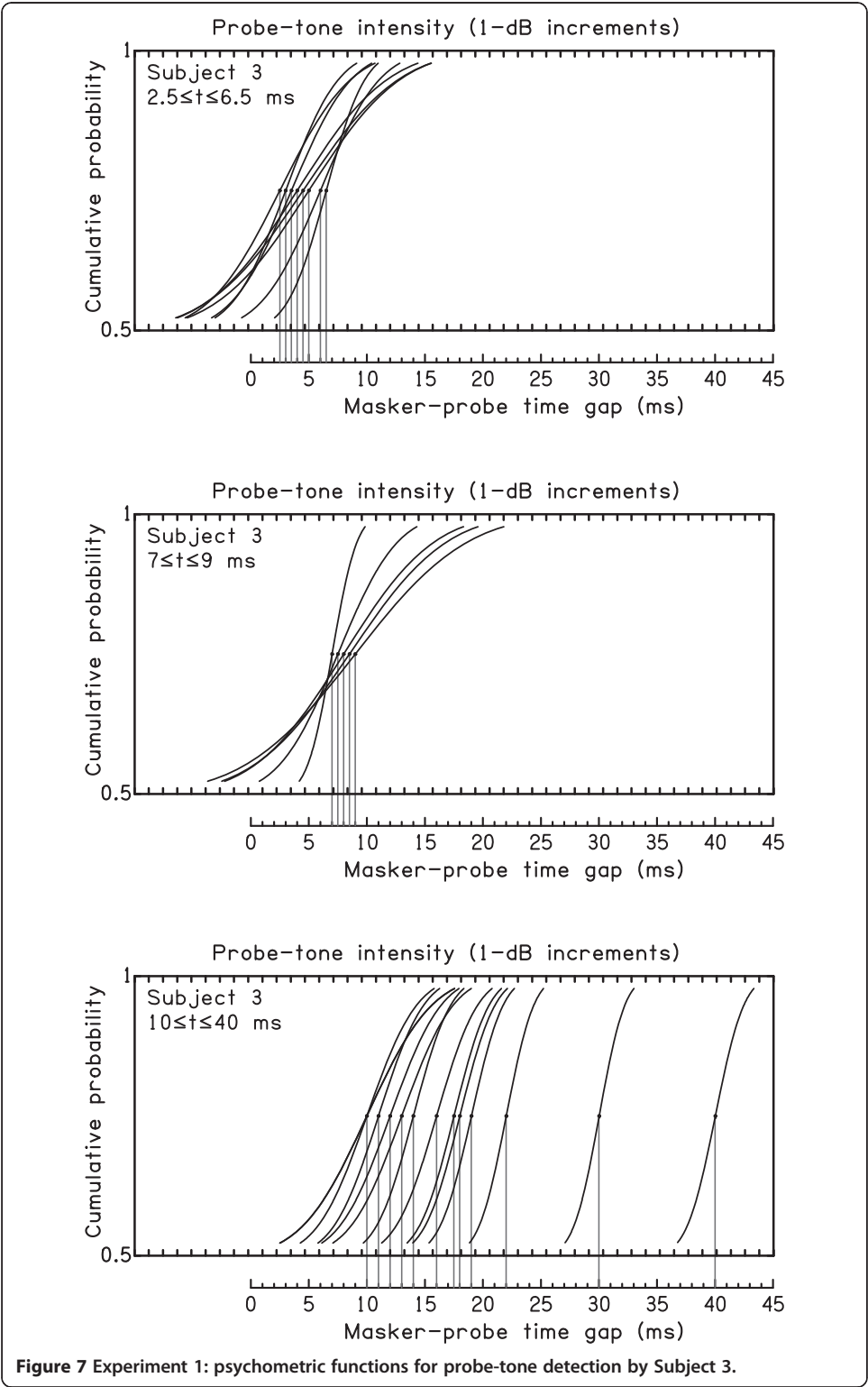


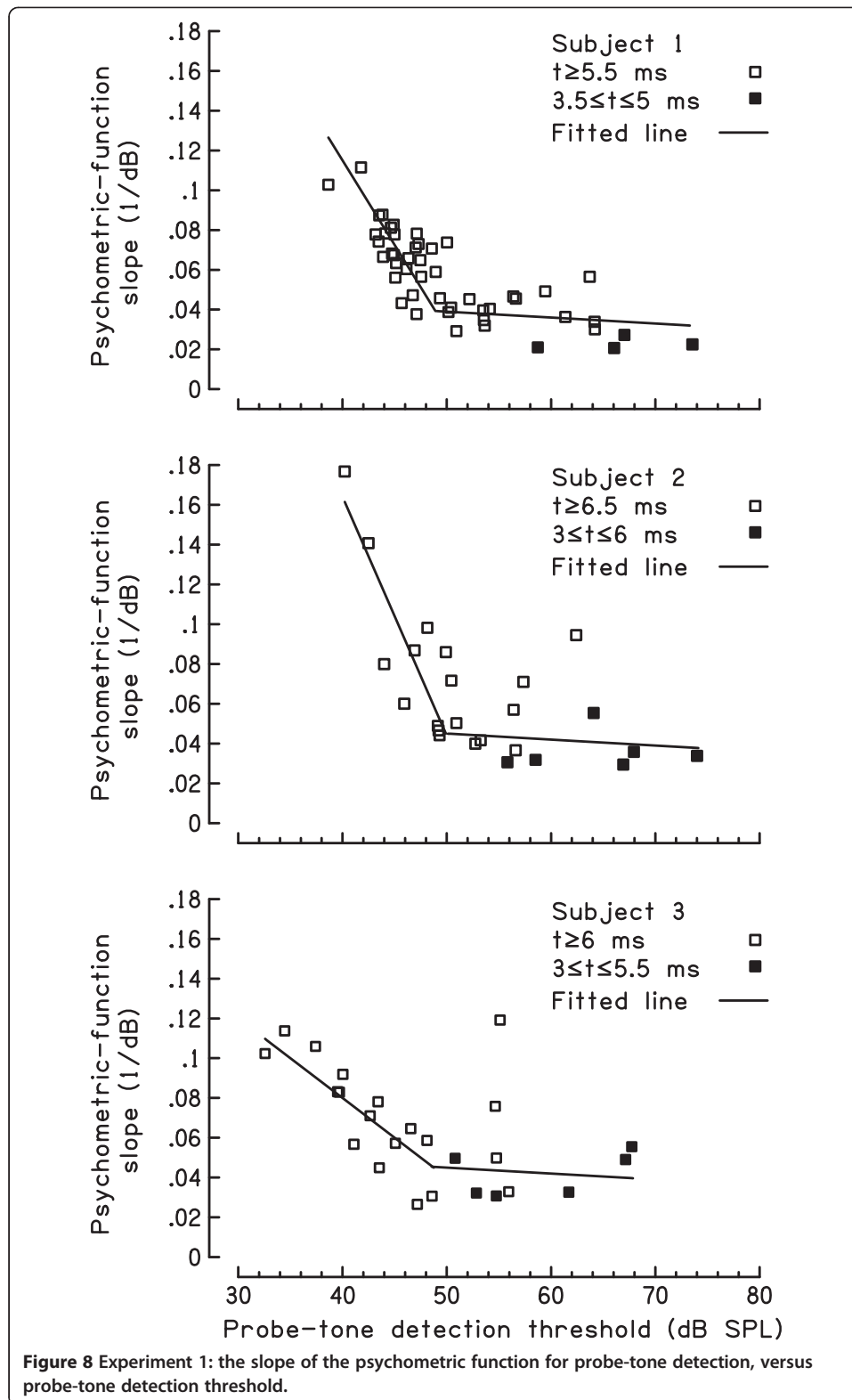
Figure 5 shows the psychometric functions for Subject 1 for $3 \leq t \leq 40$ ms. Figure 6 shows the psychometric functions for Subject 2 for $2 \leq t \leq 40$ ms. Figure 7 shows the psychometric functions for Subject 3 for $2.5 \leq t \leq 40$ ms. The psychometric functions are illustrated here without the actual data points representing the percentages-correct with which Eq. 1 was computed; those data points were so numerous that plotting them on a single graph for more than two adjacent t 's produced unreadable illustrations. Illustrations of small groups of psychometric functions, and the percentages-correct that they were fitted to, appear elsewhere [8,14].

The top and bottom horizontal scale of each panel is decibels in 1-dB increments. This is not an absolute scale in which intensity increases overall from left to right; hence the lack of numbering. Rather, each illustrated psychometric function represents its own unique range of intensities. What *does* increase from left to right, in a likewise manner for all of the midpoints of the plotted curves, is masker-probe time gap t , as shown by a vertically offset, lower horizontal scale such that a vertical line dropped from the midpoint of each psychometric function intersects the offset scale at the t for which that curve was obtained. This graphing style spreads out the psychometric functions according to masker-probe time gap rather than intensity, so that the reader can appreciate the change in the psychometric function with t . Of course, a probe-tone's detection threshold (Figure 4) – and each associated set of probe-tone intensities required to establish the psychometric function as it is shown here – *generally decrease* with increase in masker-probe time gap, although Figures 5, 6 and 7 might seem to imply the contrary.

Each subject's psychometric functions were grouped into sets representing contiguous postmasker times, the boundaries of each set being determined by the evident behavior of the probe-tone detection thresholds (after Figure 4). Namely, the sets represent (1) a steep initial threshold drop, over roughly $3 \leq t \leq 6$ ms; (2) a small threshold rise, over roughly $6.5 \leq t \leq 10$ ms; and finally (3) a gradual threshold decrease over roughly $10.5 \leq t \leq 40$ ms. For all three subjects, the forward-masker and probe-tone physically overlap roughly over $0 \leq t \leq 2.5$ ms [9]; there, the subjects perform differently than at later t , as evident by psychometric functions (omitted here for simplicity) which are steeper than those in the first panel of each of Figures 5, 6 and 7.

Recall now that, according to the Schairer et al. [4] model of the cochlear nonlinearity as two conjoined line segments, the slope of the psychometric function should take on just two values: one for low probe-tone detection thresholds, and one for moderate probe-tone detection thresholds. But Figures 5, 6 and 7 reveal more than just two unique slopes. The Schairer et al. [4] notion is therefore a simplification at best, if we wish to maintain the underlying hypothesis of Gaussian distributions of input and of output. A more sophisticated model of the cochlear nonlinearity might include what is actually observed of it in Guinea pigs and chinchillas, namely, a monotonically declining slope with increasing probe-tone intensity [2,3]. Hence, if Schairer et al. [4] are correct at least about probe-tone detection being limited by a fixed output distribution, then altogether the slopes of the psychometric functions of Experiment 1 should rise with increasing t .

Figure 8 shows the dependence of psychometric-function slope upon probe-tone detection threshold. Psychometric-function slope is expressed in units of probability points per dB ("1/dB" on the graphs), a natural-seeming unit which is 1/100 the unit of "percentage-points/dB" used by Schairer et al. [4] and by Schairer et al. [5].



Numerous attempts were made to fit equations to the open squares in Figure 8, or to the open squares plus the closed squares, without finding a visually pleasing fit^c. Therefore, the data sets in Figure 8 were described using the kind of simple approximation

used by Schairer et al. [4] for the cochlear nonlinearity. That is, a pair of conjoined line segments of form $s(x) = a x + b$ were fitted separately for each subject, where s here is *psychometric-function* slope (not to be confused with the slope of the cochlear nonlinearity itself) and x is probe-tone detection threshold in dB SPL. For Subjects 1, 2, and 3, respectively, the fitted parameters $\{a, b\}$ for the steeper line of each pair were $\{-0.0085, 0.455\}$, $\{-0.012, 0.644\}$, and $\{-0.004, 0.24\}$, and the fitted parameters for the shallower line of each pair were $\{-0.0003, 0.054\}$, $\{-0.0003, 0.06\}$, and $\{-0.0003, 0.06\}$. For Subject 1, the elbow in the conjoined lines occurs at 49.04 dB SPL (i.e., 21.2 dB above absolute probe-tone detection threshold); for Subject 2, at 50.11 dB SPL (i.e., 24.8 dB above absolute probe-tone detection threshold); and for Subject 3, at 48.55 dB SPL (i.e., 25.95 dB above absolute probe-tone detection threshold). That is, a bend is predicted in the cochlear mechanical nonlinearity at roughly 21–25 dB above the nonlinearity's minimum value, in contrast to the 30 dB above-minimum (see Figure 3) proposed by Schairer et al. [4].

Method and results of experiment 2: growth of forward-masking with increasing forward-masker intensity

Method, probe-tone detection thresholds, and psychometric functions

In Experiment 2, the durations, frequencies, and shapes of the forward-masker and the probe-tone were the same as in Experiment 1. The primary difference was that the masker-probe time gap was held constant, at 3 ms, just beyond the range of the physical overlap of forward-masker and probe-tone. Such overlap could have allowed listeners to perform discrimination rather than detection.

The 2-kHz probe-tone's detection threshold in quiet was obtained for Subject 2A; it was 26.7 dB SPL. (The quiet threshold for Subject 1A had been obtained in Experiment 1 and was 27.8 dB SPL.) Subsequently, psychometric functions for probe-tone detection were obtained for different intensities of the forward-masker. Forward-masker intensities of 30-90 dB SPL in 5-dB steps were used for Subject 1A. Subject 2A did not have as much time available as Subject 1A, and therefore ran at fewer forward-masker intensities, namely, 30, 40, 50, 55, 60, 70, 75, and 85 dB SPL. Once a probe-tone detection threshold had been established for a given forward-masker intensity, the next forward-masker's intensity was chosen, in a pseudo-random fashion.

Figure 9 shows probe-tone detection thresholds versus forward-masker intensity. The vertical error bars are the 95% confidence intervals, established through Probit Analysis. Their small size reflects better precision than any similar published work. The thresholds rise monotonically, as generally seen in the literature and as found for other stimulus-and-masker combinations by Schairer et al. [4, Figures. 2 & 6] and by Schairer et al. [5, Figure 2].

Figure 10 shows the psychometric functions for Experiment 2 (smooth curves), generated by fitting Eq. 1 to the empirical percentages-correct using Probit Analysis. Each psychometric function is labeled by the intensity of the forward-masker used. The psychometric functions generally widen with increase in forward-masker intensity, and hence with increase in probe-tone detection threshold. There was one notable exception: the psychometric function for Subject 1A for the 70 dB SPL (open squares) forward-masker is anomalously steep.

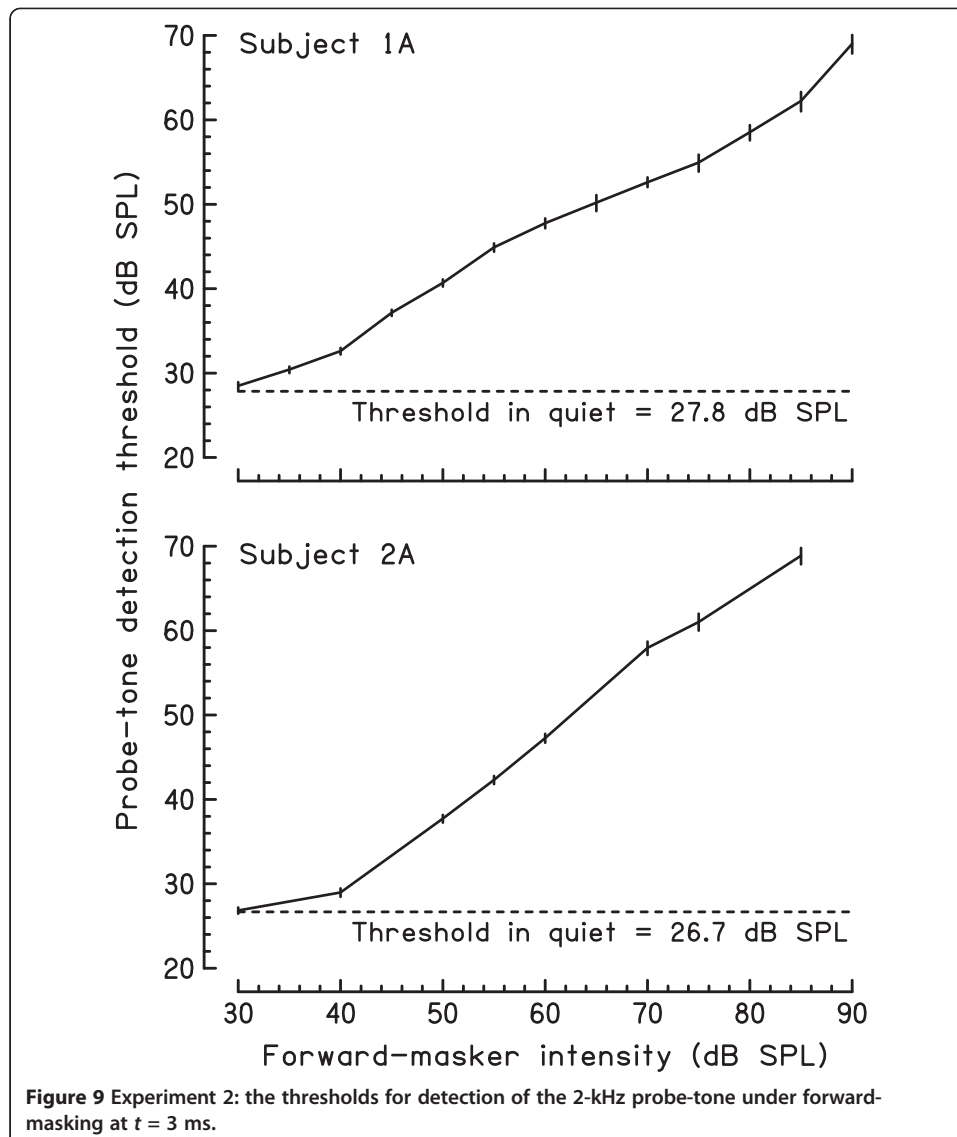
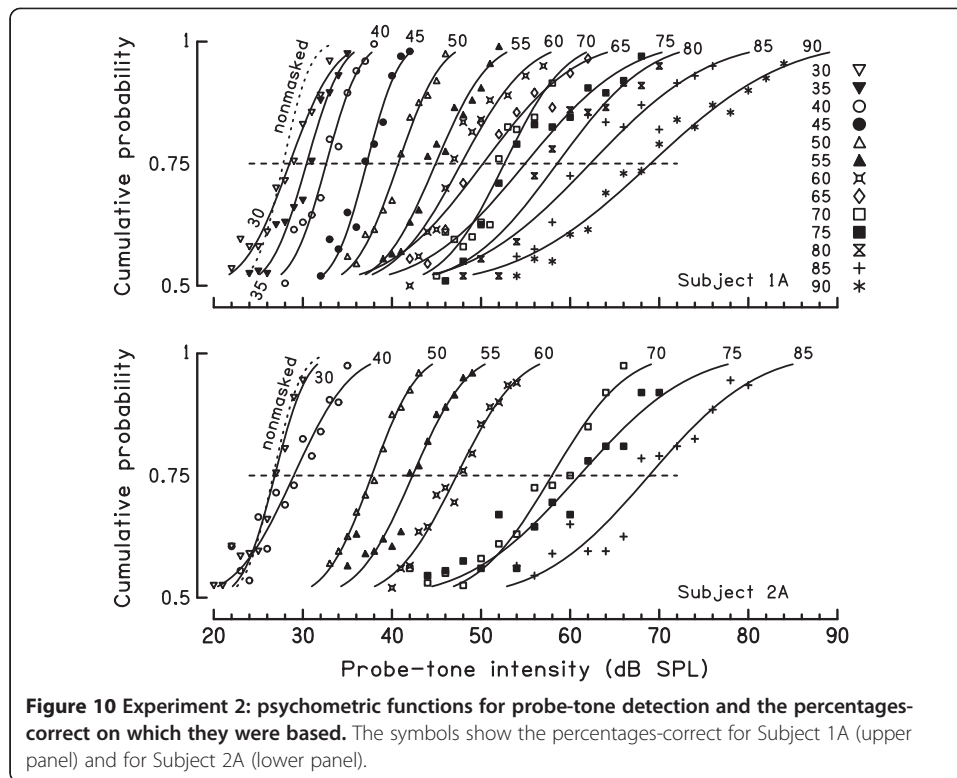


Figure 11 shows the slopes of the psychometric functions, plotted versus the respective probe-tone detection thresholds. Let the psychometric-function slope be s and let the probe-tone detection threshold be x . For Subject 1A, $s(x) = 53x^{-1.9}$. For Subject 2A, $s(x) = 2.4x^{-1.12}$. Hence, if psychometric functions for forward-masked probe-tone detection are indeed determined by a nonlinear cochlear input/output relation and multiplicative internal noise, as hypothesized by Schairer et al. [4], then the smooth decline in psychometric-function slope seen for both subjects in Figure 11 implies a cochlear nonlinearity that becomes increasingly compressive (i.e., shallow) with increasing probe-tone intensity. The difference in the exponents in $s(x)$ may reflect differences in the cochlear nonlinearity across listeners.

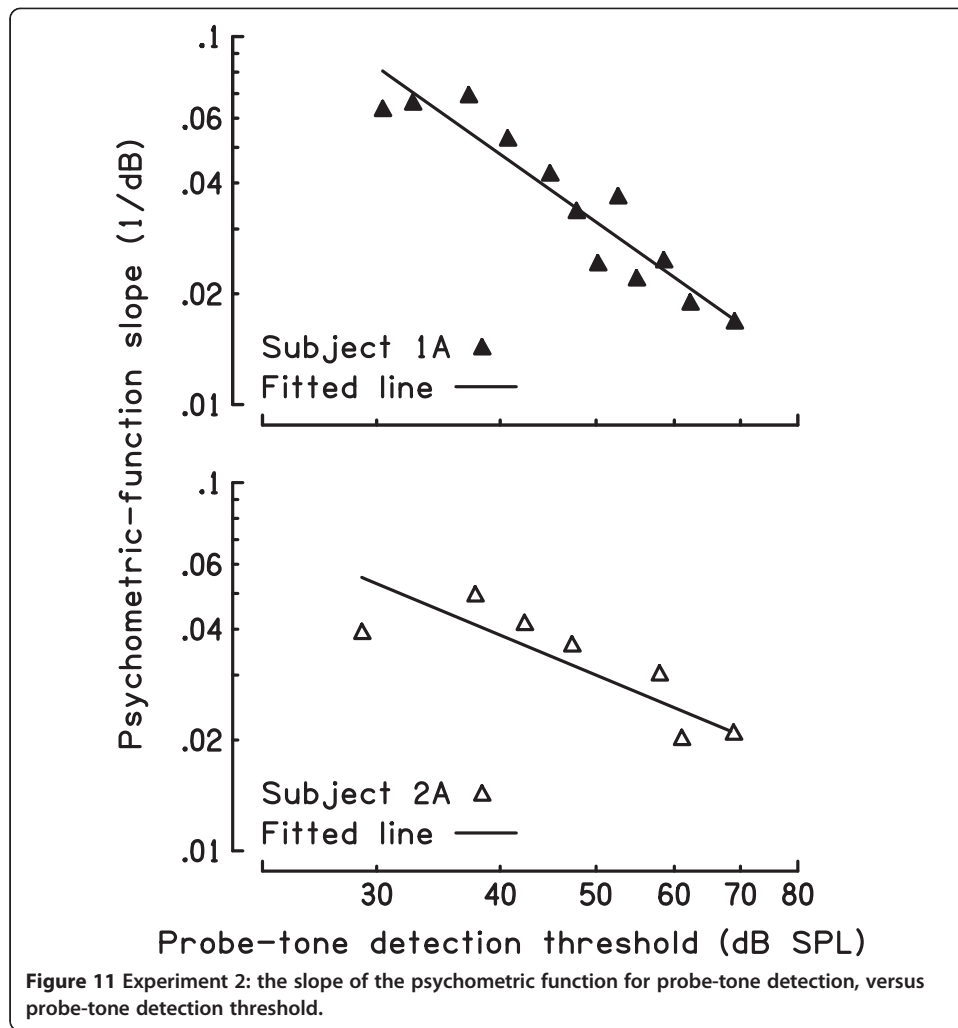
Comparison data: Schairer et al. [5]

Schairer et al. [5, Figure 4] had obtained psychometric functions for five subjects for the detection of 10-ms tones of 0.25 kHz or 4 kHz, forward-masked at a masker-probe



time gap of 10 ms (forward-masker offset to probe-tone onset) by same-frequency 200-ms forward-maskers of various intensities (their “Experiment 1”). It proved appropriate and instructive to re-examine the slopes of those psychometric functions in the same manner as the slopes from the present Experiment 2. Figure 12 shows the Schairer et al. [5] psychometric-function slopes pooled over all five of their subjects, each plotted point thus representing a single subject and a single forward-masker intensity. For the sake of comparison to the present psychometric-function slopes, the Schairer et al. [5] units of slope have been multiplied by 1/100. For 0.25 kHz we obtain $s(x) = 3.7x^{-1.65}$, and for 4 kHz we obtain $s(x) = 2.65x^{-1.65}$. Note well that the coincidental common exponent of the two latter equations is not incompatible with the exponents for the 2 kHz probe-tone of Experiment 2. Once again, within the initial Schairer et al. [4] model assumptions, the cochlear nonlinearity implied from [5, Figure 4] is increasingly compressive over the entire course of the employed probe-tone intensities.

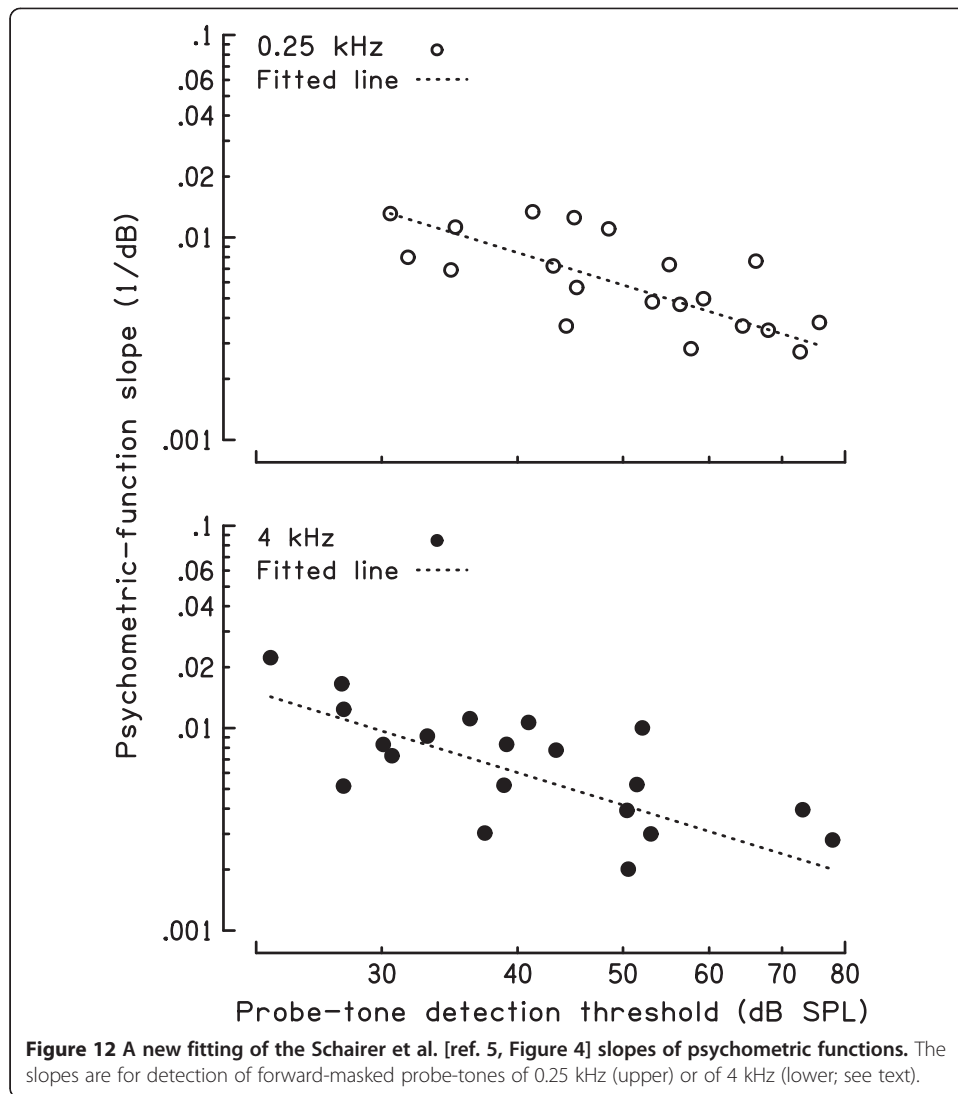
Consider, however, the sheer magnitudes of the psychometric-function slopes. Figure 13 shows the lines produced by the equations fitted to psychometric-function slopes as a function of probe-tone detection threshold (Figures 11 and 12). That is, the upper, solid straight lines in Figure 13 represent the equations $s(x) = 53x^{-1.9}$ for Subject 1A and $s(x) = 2.4x^{-1.12}$ for Subject 2A. Also shown are dotted straight lines which respectively represent the equation $s(x) = 3.7x^{-1.65}$ for the 0.25 kHz probe-tone detection thresholds of Schairer et al. [5] and the equation $s(x) = 2.65x^{-1.65}$ for the 4 kHz probe-tone detection thresholds of Schairer et al. [5]. The latter equations sit lower than those for the probe-tone detection thresholds of the present Experiment 2, as indicated by the horizontal dashed line, which is merely a visual aid. This difference presumably



reflects Schairer et al.'s [5] underestimation of psychometric-function slopes, as also evident in Schairer et al. [4] (more on this in the Discussion). Of course, the between-study differences seen in Figure 13 could reflect a less compressive cochlear mechanical nonlinearity (that is, a *steeper* nonlinearity in Figure 3) at the 2 kHz place on the basilar membrane than at 0.25 kHz or at 4 kHz, hence leading to steeper psychometric functions for the 2 kHz probe-tone according to the Schairer et al. [4] model. But Figure 13 shows that the psychometric-function slopes for the 70 dB SPL 2-kHz probe-tones of Experiment 2 were similar to those for the probe-tones near or below 30 dB SPL in Schairer et al. [5] – which would have required a truly profound difference in compression, even given the differences in the energies of the tones.

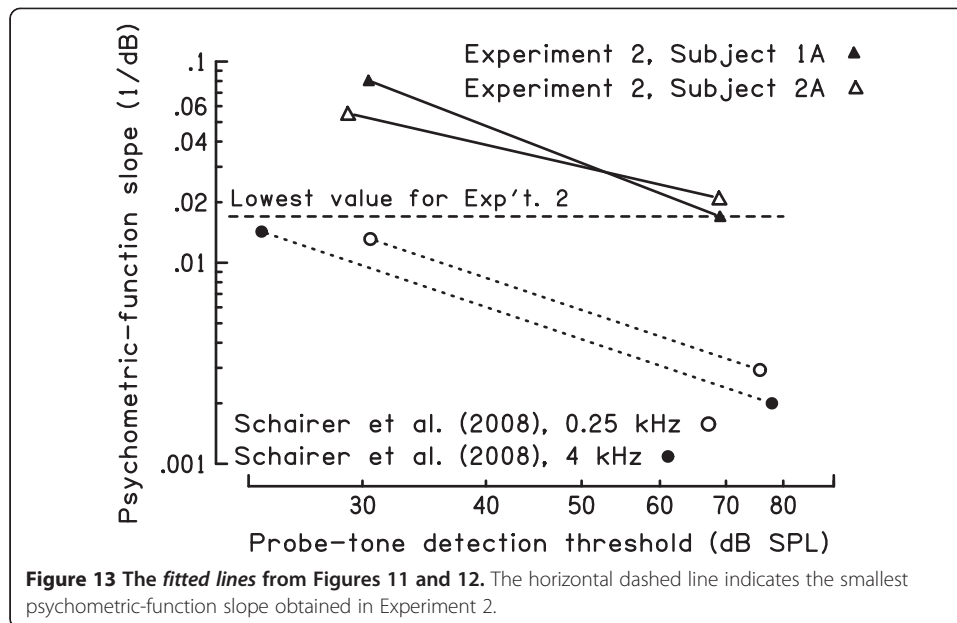
Comparison data: Schairer et al. [4]

Schairer et al. [4] had obtained psychometric functions for detection of a 10-ms 4-kHz tone which followed a 200-ms forward-masker at either of 0, 10, or 30 ms (forward-masker offset to probe-tone onset). The forward-masker's frequency was either 4 kHz



or 2.4 kHz. The psychometric-function slopes for those probe-tone detection thresholds (for subjects labeled “S4” to “S9” in Experiment 2 of Schairer et al. [4]) were reanalyzed here for the 4 kHz forward-masker in a manner similar to what was done above for psychometric-function slopes of Schairer et al. [5]. First, unlike the original illustration made by Schairer et al. [4, Figure 8], which grouped the discovered psychometric-function slopes subject-by-subject, all of those slopes were instead pooled across six of their seven listeners into scatterplots of psychometric-function slope versus probe-tone detection threshold, according to masker-probe time gap. (A seventh listener, originally labeled “S10”, produced unusual data, which were excluded here.) Figure 14 shows the replots of the Schairer et al. [4, Figure 8] data.

Using s once again for psychometric-function slope and x once again for probe-tone detection threshold in dB SPL, the fitted equations for the 4 kHz forward-masker of Schairer et al. [4] and for respective time gaps of 0, 10, and 30 ms were $s(x) = 1,480x^{-3.05}$, $s(x) = 2,800x^{-3.5}$, and $s(x) = 800,000x^{-5.2}$. The respective ranges of all of the probe-tone detection thresholds were 28-88 dB SPL, 28-70 dB SPL, and 25-56 dB SPL.



Psychometric-function slopes: Schairer et al. [4] versus Schairer et al. [5]

Schairer et al. [5], like Schairer et al. [4] earlier, had obtained slopes of psychometric functions for the detection of 4-kHz 10-ms tones forward-masked at a 10-ms time gap by same-frequency 200-ms forward-maskers. For Schairer et al. [5], $s(x) = 2.65x^{-1.65}$ (Figure 12, lower panel), in contrast to $s(x) = 2,800x^{-3.5}$ for Schairer et al. [4] (Figure 14, center panel). That is, the same experimental conditions in the same laboratory resulted in very different power-function exponents when slopes were pooled over different groups of listeners, despite a similar “spread” of slopes (roughly 0.002-0.025). A possible factor for this difference is profound intersubject variability, which has the potential to produce greater discrepancies as subject groups get smaller.

Analysis (1): extending the Schairer et al. [4] model to include the cochlear nonlinearity's average rate-of-change with intensity

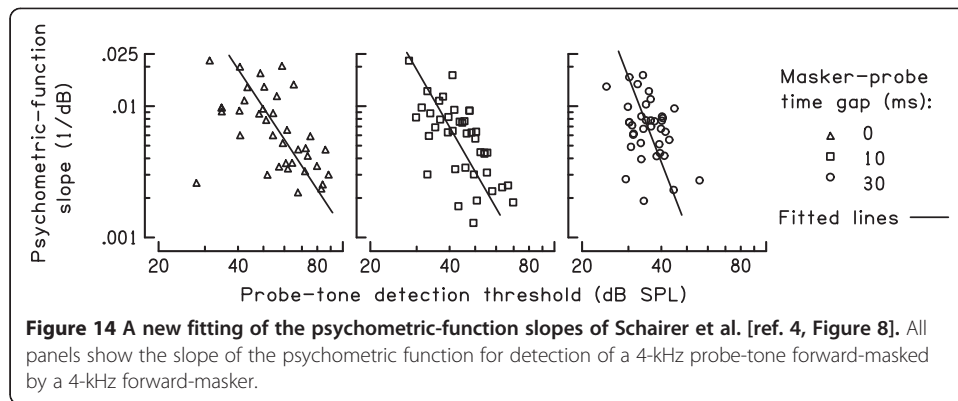
A. A new assumption and its consequences: the relation of the Gaussian-shaped “input distribution” to the “width” of the psychometric function

Under the Schairer et al. [4] model, the slope of the cochlear nonlinearity relates to the width of the psychometric function for probe-tone detection. Precisely how was never expressed in equations, however. Equations will now be provided. They depend upon a new assumption, as follows, one which was not stated in the Schairer et al. [4] model.

For forward-masked probe-tone intensity x in decibels SPL, consider a Gaussian probability density function $p(x)$, having units of probability per decibel. Suppose that $p(x)$ is the Gaussian-shaped “input distribution” of the Schairer et al. [4] model

(Figure 3). Its integral is, $P(x) = \int_{-\infty}^x p(x) dx$, the (unitless) cumulative probability as a

function of x in decibels SPL. It is now assumed that when percentage-correct scores from two-interval two-alternative forced-choice experiments such as Experiments 1 and 2 are



divided by 100, then $P(x)$ is given by Eq. 1. Those $P(x)$'s, used in Figures 5, 6 and 7 and in Figure 10, are each characterized by some mean value μ in dB SPL and by some standard deviation $\sigma(x)$ in dB. For some positive integer n , an even-numbered positive multiple of σ , call it $2n\sigma(x)$ in dB, can be reasonably considered as a convenient measure of the width of the ogive, as long as a consistent n is used.

The two endpoints (probe-tone intensities) which characterize the “width” of a psychometric function define points on the cochlear nonlinearity such that the latter's average slope, call it \bar{f} , can be measured between said points. Narrowing that span in equal decrements from each endpoint causes the average slope of the cochlear nonlinearity to approach the instantaneous slope at the intensity corresponding to the centroid of the psychometric function – the intensity taken as the probe-tone detection threshold, μ . Figure 15 illustrates these concepts.

The span $2n\sigma(x)$ is the denominator, in decibels, of the ratio needed to calculate the cochlear nonlinearity's average slope \bar{f} over $(\mu - n\sigma)$ to $(\mu + n\sigma)$. Of course, a numerator is also needed – the respective change in *output*. That change, in the Schairer et al. [4] model, is assumed to be a constant number of dB, an unknown that will be called κ here. Making κ a constant allows the denominator of \bar{f} to be set simply to $\sigma(x)$ alone, because any multipliers of $\sigma(x)$ will be combined with κ to give a new unknown, k :

$$\bar{f}(x) = \frac{k}{2n\sigma(x)} = \frac{k}{\sigma(x)} \text{ dB/dB} \quad (3)$$

Recall from above that for Experiments 1 and 2, psychometric-function slope s in 1/dB relates to psychometric-function σ in dB as $s = 1/(\sigma\sqrt{8\pi})$. Recall also that s is empirically quantified above in equations of form $s = s(x)$. Altogether, *the average slope of the cochlear nonlinearity* over the width of the psychometric function having its centroid at $x = \mu$ is

$$\bar{f}(x) = \frac{k}{\sigma(x)} = K \cdot s(x) \text{ dB/dB, where } K = k\sqrt{8\pi} \quad (4)$$

How accurate is Eq. 4? Let us momentarily call the nonlinearity itself $F(x)$. For the probe-tone detection threshold $x=\mu$ (Figure 15), the actual average slope of the

Overall, then, Eq. 4 is accurate within the Schairer et al. [4] two-line-segment model for any psychometric function whose boundaries of declared “width” puts it *entirely within* either branch of those two line segments. For widths that span the “hinge” where the two line segments meet, however, Eq. 4 may fare poorly. Ironically, it should not fare as poorly for the kind of realistic gradual changes in cochlear-nonlinearity slope which are shown in Figure 18 (below), but it will still not fare as well as it would

within the Schairer et al. [4] two-branch model, for psychometric functions whose “width” falls entirely within either branch.

Note again Eq. 4, which shows that $\bar{f}(x)$, the predicted average slope of the cochlear nonlinearity, depends upon $s(x)$, the psychometric-function slope s as a function of the probe-tone detection threshold x . Examples of that $s(x)$ are the paired line segments of Figure 8. When substituted into Eq. 4, they yield for Subject 1 of Experiment 1 two solutions, each for a different intensity range:

$$\begin{aligned}\bar{f}(x) &= K \cdot s_1(x) \\ &= K \cdot (-0.0085x + 0.455) \text{ dB/dB}, \quad 38.6 \leq x \leq 48.9 \text{ dB SPL}\end{aligned} \quad (6a)$$

$$\begin{aligned}\bar{f}(x) &= K \cdot s_2(x) \\ &= K \cdot (-0.0003x + 0.054) \text{ dB/dB}, \quad 48.9 \leq x \leq 73.6 \text{ dB SPL}\end{aligned} \quad (6b)$$

and likewise for Subject 2 of Experiment 1

$$\bar{f}(x) = K \cdot s_1(x) = K \cdot (-0.012x + 0.644) \text{ dB/dB}, \quad 40.2 \leq x \leq 49.9 \text{ dB SPL} \quad (7a)$$

$$\bar{f}(x) = K \cdot s_2(x) = K \cdot (-0.0003x + 0.06) \text{ dB/dB}, \quad 49.9 \leq x \leq 74.0 \text{ dB SPL} \quad (7b)$$

and likewise for Subject 3 of Experiment 1

$$\bar{f}(x) = K \cdot s_1(x) = K \cdot (-0.004x + 0.24) \text{ dB/dB}, \quad 32.6 \leq x \leq 48.6 \text{ dB SPL} \quad (8a)$$

$$\bar{f}(x) = K \cdot s_2(x) = K \cdot (-0.0003x + 0.06) \text{ dB/dB}, \quad 48.6 \leq x \leq 67.7 \text{ dB SPL} \quad (8b)$$

Consider a contrasting set of $s(x)$'s, namely the power functions of Figure 11. When they are substituted into Eq. 4, we obtain for Subject 1A of Experiment 2

$$\bar{f}(x) = K \cdot 53x^{-1.9} \text{ dB/dB}, \quad 30.4 \leq x \leq 69.0 \text{ dB SPL} \quad (9)$$

and for Subject 2A of Experiment 2

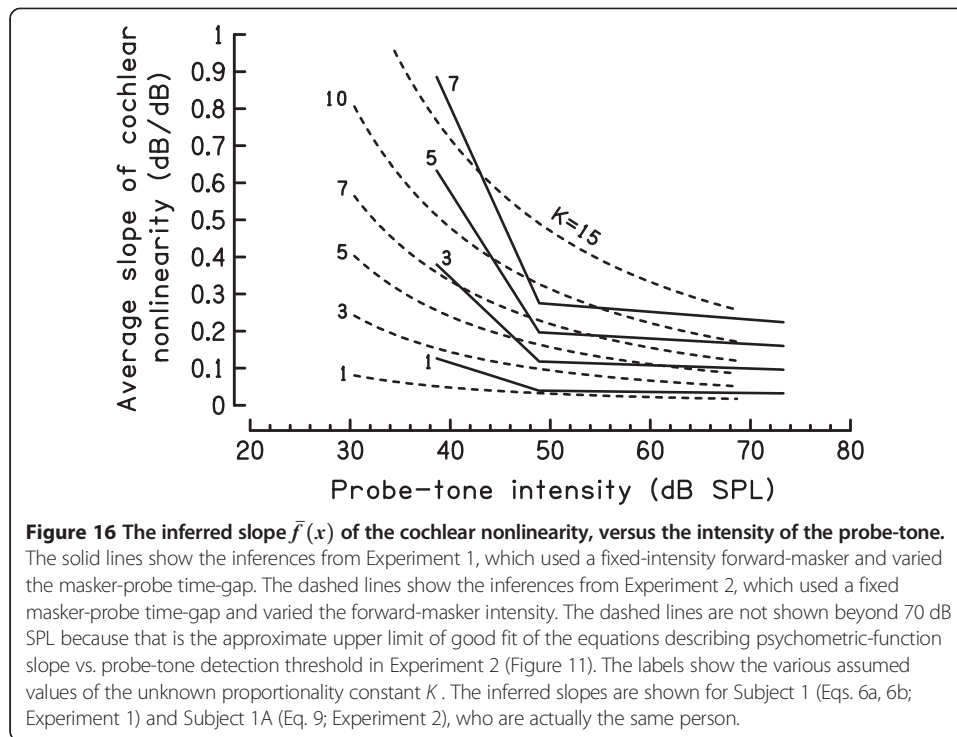
$$\bar{f}(x) = K \cdot 2.4x^{-1.12} \text{ dB/dB}, \quad 29 \leq x \leq 68.9 \text{ dB SPL} \quad (10)$$

Curves can be plotted for $\bar{f}(x)$ for Eqs. 6, 7, 8, 9 and 10 if values of the unknown, K , are chosen. Figure 16 shows such curves for Subjects 1 and 1A, who are the same person. Eqs. 6, 7, 8, 9 and 10 give similar magnitudes of $\bar{f}(x)$ as Eqs. 9 and 10 when the latter are assigned values of K that are roughly double those assigned to Eqs. 6, 7 and 8.

Unfortunately, K cannot be specified *a priori*. Therefore which of the curves of Figure 16 are realistic can only be inferred by comparison to curves of the slope of the cochlear nonlinearity derived empirically, which are available from animals.

The cochlear nonlinearity's rate-of-change with intensity in animals

Empirical properties of the cochlear nonlinearity as a function of stimulus intensity have been measured in several species of small mammals, placed under anesthesia (e.g., [15-17]). Such measurements allow inference of the rate-of-change of the cochlear nonlinearity.



Nuttall and Dolan [16] obtained basilar-membrane displacement *velocities* in Guinea pigs as a function of the intensity of 150-ms pure tones. Nuttall and Dolan's velocities, expressed in meters/s, were here converted to displacements in meters, by dividing by 2π times the tone frequency in Hz (A.L. Nuttall, personal communication)^d. Those displacements were denoted D , and were expressed in decibels by taking 10 times their logarithm to base 10. Decibel displacements were then plotted versus the respective evoking tone intensities x in dB SPL. Three such plots were made, using those measurements which employed the greatest available number of evoking tone intensities (which was not necessarily constant from animal to animal or from tone frequency to tone frequency), allowing better discrimination of the best-fitting equation.

Decibels of displacement was fitted as a function of intensity x according to

$$D(x) = a \cdot (x + b)^c - \frac{d}{x} \quad \text{dB} \quad (11)$$

Equation 11 was chosen out of numerous candidate equations because of its fit, which was visually pleasing for all but the rightmost four or five data points of each plot.^e The displacements and the curves fitted to them, along with the fitted parameters $\{a, b, c, d\}$, are not shown here for the sake of brevity, but are available from the author on demand. The slope of the cochlear input-output response as a function of x is given by the derivatives $dD(x)/dx$, where from Eq. 11,

$$\frac{dD(x)}{dx} = ca \cdot (x + b)^{c-1} + \frac{d}{x^2} \quad \text{dB/dB} \quad (12)$$

The fitted parameters of Eq. 11 are substituted into Eq. 12, case-by-case.

Basilar-membrane peak displacements are available for 30-ms pure tones in chinchillas [15]. Four plots were made, from measurements using greater numbers of evoking tone intensities. The displacements were fitted to

$$D(x) = a \cdot (x + b)^c - \left(\frac{d}{x + 6} \right) \text{ dB} \quad (13)$$

where the 6 in the denominator of the second term is necessary to avoid a zero denominator when $x \leq 0$ (which happened for one particular plot for which x fell to -5 dB SPL). Again, the displacements and the curves fitted to them, along with the fitted parameters $\{a, b, c, d\}$, are not shown here for the sake of brevity, but are available from the author on demand. To get the slope of the cochlear nonlinearity, the fitted parameters of Eq. 13 are substituted case-by-case into

$$\frac{dD(x)}{dx} = ca \cdot (x + b)^{c-1} + \frac{d}{(x + 6)^2} \text{ dB/dB} \quad (14)$$

Finally, no curvefitting is needed for the chinchillas of Ruggero et al. [17, Figure 3]; Ruggero et al. actually provided slopes of their cochlear nonlinearities.

Analysis (2): comparing the cochlear nonlinearity's rate-of-change in man to that in animals

Figures 17, 18, 19, 20 and 21 compare the *slopes* of the nonlinearity for humans to those of animals. The latter slopes have a general downwards trend, as expected for cochlear nonlinearities that become shallower (i.e., increasingly compressive) with increasing sound-pressure-level. Also, the slopes inferred from Nuttall and Dolan ([16]; upper panels of Figures 17, 18, 19, 20 and 21) and from Rhode and Recio ([15]; middle panels of Figures 17, 18, 19, 20 and 21) are of the same *magnitudes* as those that were actually measured by Ruggero et al. ([17]; lower panels of Figures 17, 18, 19, 20 and 21), a reassuring consistency.

Figure 17 compares the inferred slope of the cochlear nonlinearity for a human to those of Guinea pigs and of chinchillas. For the sake of comparing smooth curves to smooth curves, it was initially most convenient to use smooth equation-generated curves for humans (those arising from Experiment 2), rather than angled, bi-segmented functions (those arising from Experiment 1). Specifically, nonlinearity slopes for Subject 1A (the dashed lines of the upper panel of Figure 16) are compared to the slopes generated for animals by Eq. 13 (upper panel of Figure 17) or Eq. 15 (middle panel of Figure 17), or to actual measured slopes for animals (connected points in lower panel).

Importantly, the curves inferred for the slope of the cochlear nonlinearity in animals (and the lower-panel data) all had to be shifted rightwards, i.e., to higher SPLs, in order to make them comparable to those inferred from Experiments 1 and 2. Why? The auditory periphery in humans does not grossly differ anatomically or physiologically from that in small mammals (e.g., [18]), although the frequency range of hearing, for example, can of course differ. The needed rightwards shift may be due to differences in the probe-tone stimuli themselves, as follows. Excluding the brief durations for up-and-down-ramping to maximum sine-wave amplitude, the pure tones used by Nuttall and Dolan [16] were 150 ms long, those of Rhode and Recio [15] were 30 ms long, and those of Ruggero et al. [17] were up to 25 ms in duration. The present Gaussian-

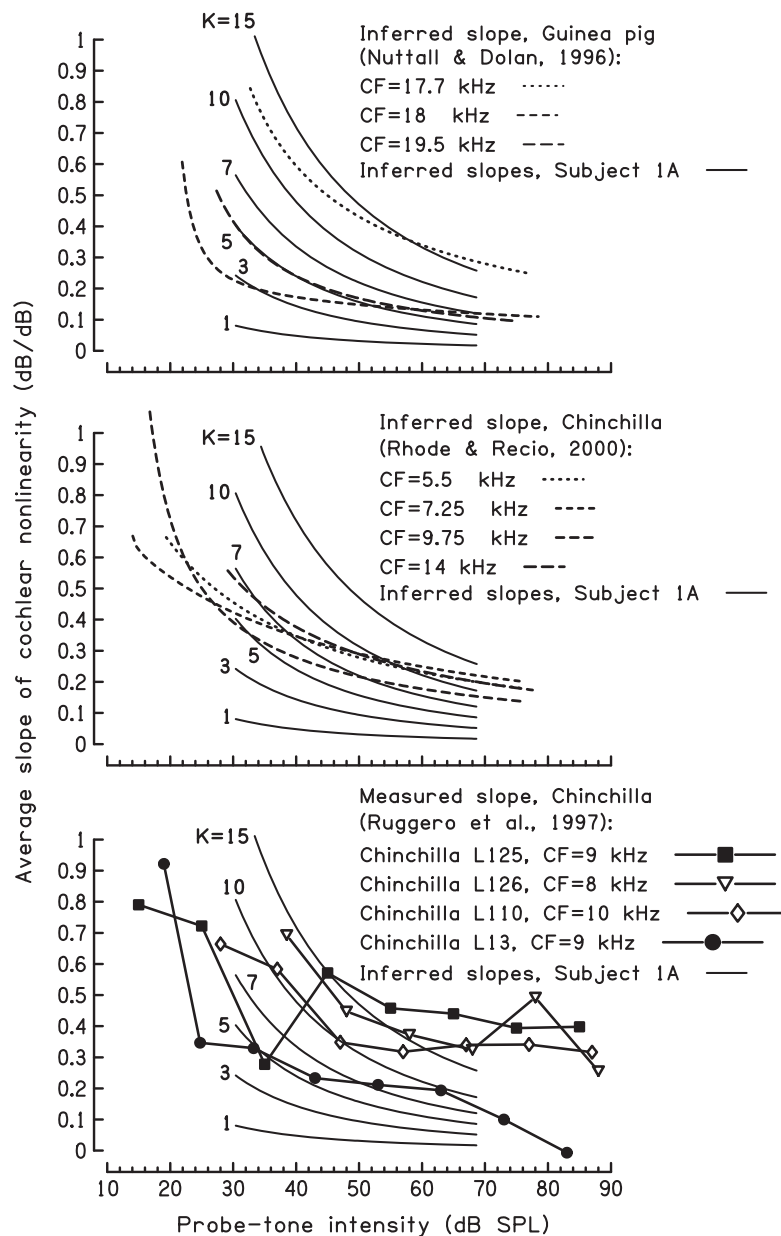


Figure 17 Comparison of inferred slopes of the cochlear nonlinearity: Subject 1A versus animals.

The inferred slopes of the cochlear nonlinearity for Subject 1A (continuous curves, same as the broken curves from the upper panel of Figure 16) are overlaid upon dashed curves (Eq. 12 for upper panel, Eq. 14 for middle panel) or data (for lower panel) for the nonlinearity slopes from Guinea pigs or from chinchillas. The solid curves here are the same in the upper, middle, and lower panels, and represent different values of the unknown proportionality constant K for Subject 1A. Rightward shifts, in decibels, were applied to the physiologically-derived curves or plots (see text); those shifts, and their allowed ranges in decibels are as follows: (upper panel) CF=17.7 kHz, shift=15.54 dB, allowed range={9.47,22.48}; CF=18 kHz, shift=18.54 dB, allowed range={9.54, 22.55}; CF=19.5 kHz, shift=13.54 dB, allowed range={9.89, 22.90}; (middle panel) CF=5.5 kHz, shift=15.54 dB, allowed range={4.39, 17.40}; CF=7.25 kHz, shift=19.54 dB, allowed range={5.59, 18.60}; CF=9.75 kHz, shift=14.54 dB, allowed range={6.88, 19.89}; CF=14 kHz, shift=11.54 dB, allowed range={8.45, 21.46}; (lower panel) CF=9 kHz, shift=10 dB, allowed range={6.53, 19.54}; CF=8 kHz, shift=13 dB, allowed range={6.02, 19.03}; CF=10 kHz, shift=12 dB, allowed range={6.99,20}; CF=9 kHz, shift=8 dB, range={6.53, 19.54}.

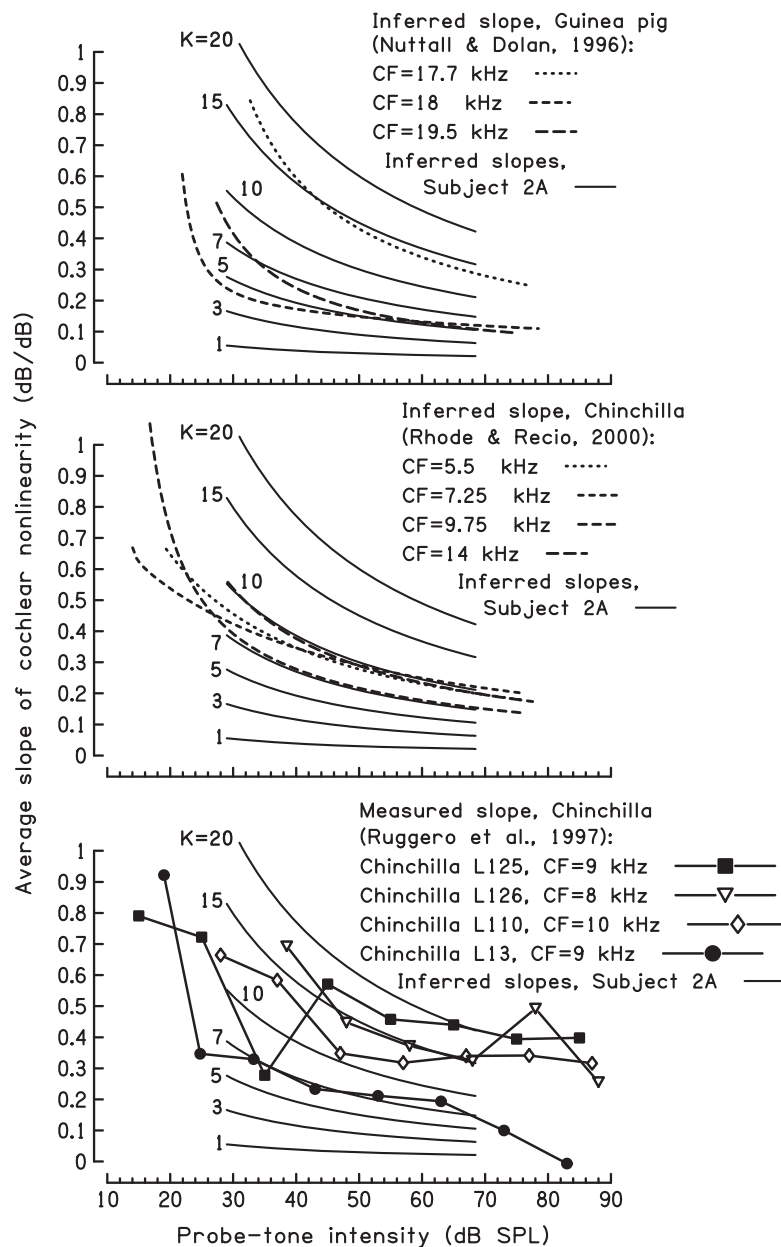


Figure 18 Comparison of inferred slopes of the cochlear nonlinearity: Subject 2A versus animals.

The inferred slopes of the cochlear nonlinearity for Subject 2A are overlaid upon dashed curves (Eq. 12 for upper panel, Eq. 14 for middle panel) or data (for lower panel) for the nonlinearity slopes from Guinea pigs or chinchillas. The solid curves here are the same in the upper, middle, and lower panels, and represent different values of the unknown proportionality constant K for Subject 2A. Rightward shifts, in decibels, were applied to the physiological curves or plots, as described in the caption to Figure 17.

shaped probe-tones had an equivalent rectangular duration – the duration of a comparable tone of constant amplitude – of approximately $0.5\sqrt{2\pi} = 1.25$ ms [19].

The integrated energy of a stimulus over its duration is directly proportional to its equivalent rectangular duration. If we momentarily ignore possible energy differences associated with differences of probe-tone *frequency*, then the present probe-tone had, at most, 1/120th of the total energy of the probe-tones used by Nuttall and Dolan [16], 1/24th of the total energy of those used by Rhode and Recio [15], and 1/20th of the

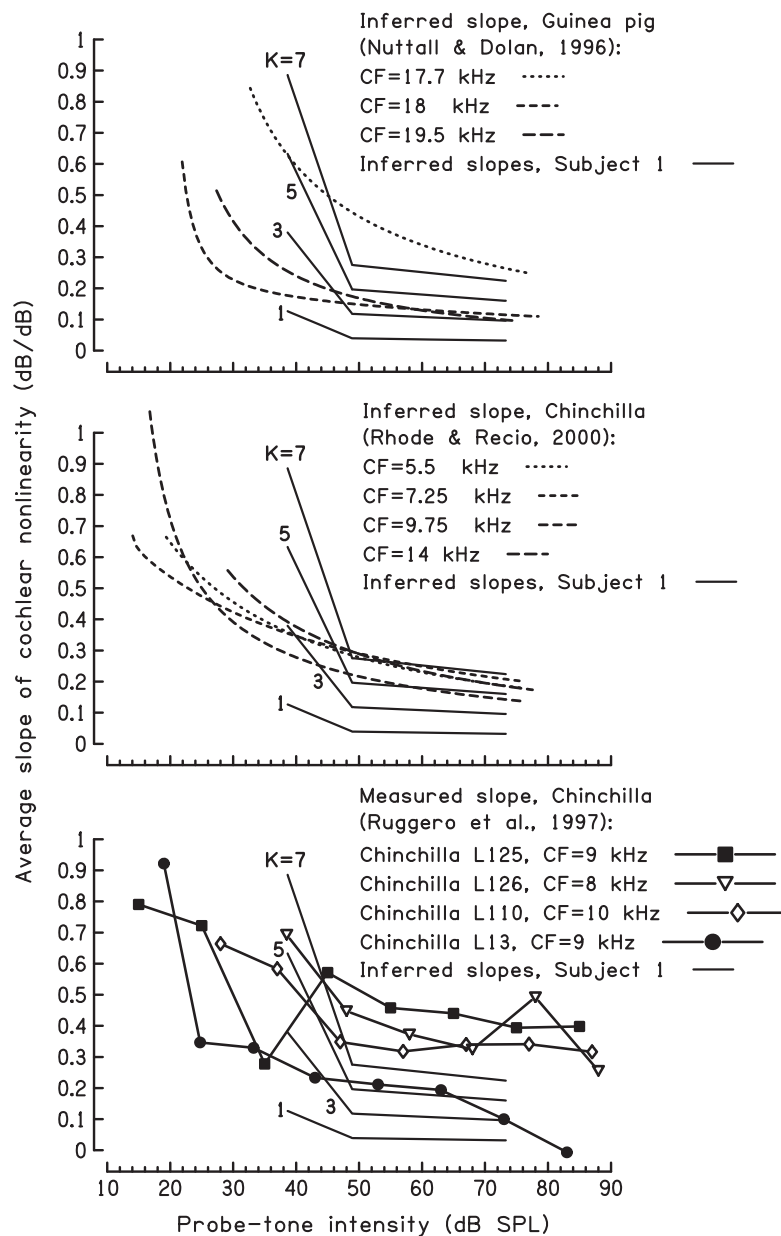


Figure 19 Comparison of inferred slopes of the cochlear nonlinearity: Subject 1 versus animals. The inferred slopes of the cochlear nonlinearity for Subject 1 (the bisegmented lines from Figure 16) are overlaid upon dashed curves (Eq. 12 for upper panel, Eq. 14 for middle panel) or data (for lower panel) for the nonlinearity slopes from Guinea pigs or chinchillas. The bisegmented lines here are the same in the upper, middle, and lower panels, and represent different values of the unknown proportionality constant K for Subject 1. Rightward shifts, in decibels, were applied to the physiological curves or plots, as described in the caption to Figure 17.

total energy of the longest tones used by Ruggero et al. [17]. Furthermore, the energy available from the present probe-tone at its 2-kHz CF place on the basilar membrane is even less than suggested by its equivalent rectangular duration, because the probe-tone's rising and falling envelope spreads its energy frequency-wise, giving it a relative spectral energy density which has a single lobe spanning 1.517-2.483 kHz at 10 dB below its maximum.

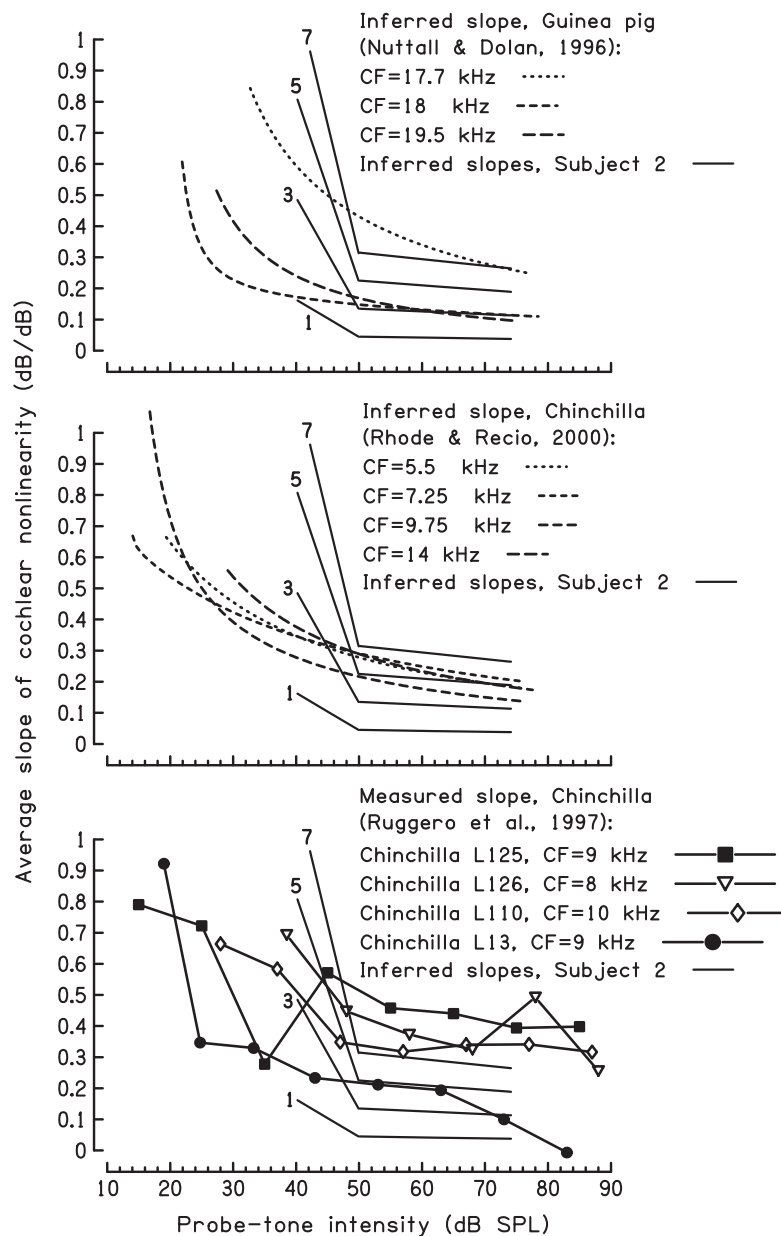


Figure 20 Comparison of inferred slopes of the cochlear nonlinearity: Subject 2 versus animals. The inferred slopes of the cochlear nonlinearity for Subject 2 are overlaid upon broken curves (Eq. 12 for upper panel, Eq. 14 for middle panel) or data (for lower panel) for the nonlinearity slopes from furred mammals. The bisegmented lines here are the same in the upper, middle, and lower panels, and represent different values of the unknown proportionality constant K for Subject 2. Rightward shifts, in decibels, were applied to the physiological curves or plots, as described in the caption to Figure 17.

It is no surprise, then, that the present probe-tone required higher stimulus intensities to activate the basilar membrane to similar degrees of responsiveness as the stimuli of the cited physiological studies; it had much less driving energy. We may compensate for that lack by moving the plots for Nuttall and Dolan [16], Rhode and Recio [15], and Ruggero et al. [17] up-intensity. But careful inspection of the original papers reveals that, in terms of maximum displacement, the Nuttall and Dolan [16] 150-ms tone was not more effective in driving the basilar membrane than was the Ruggero et al. [17]

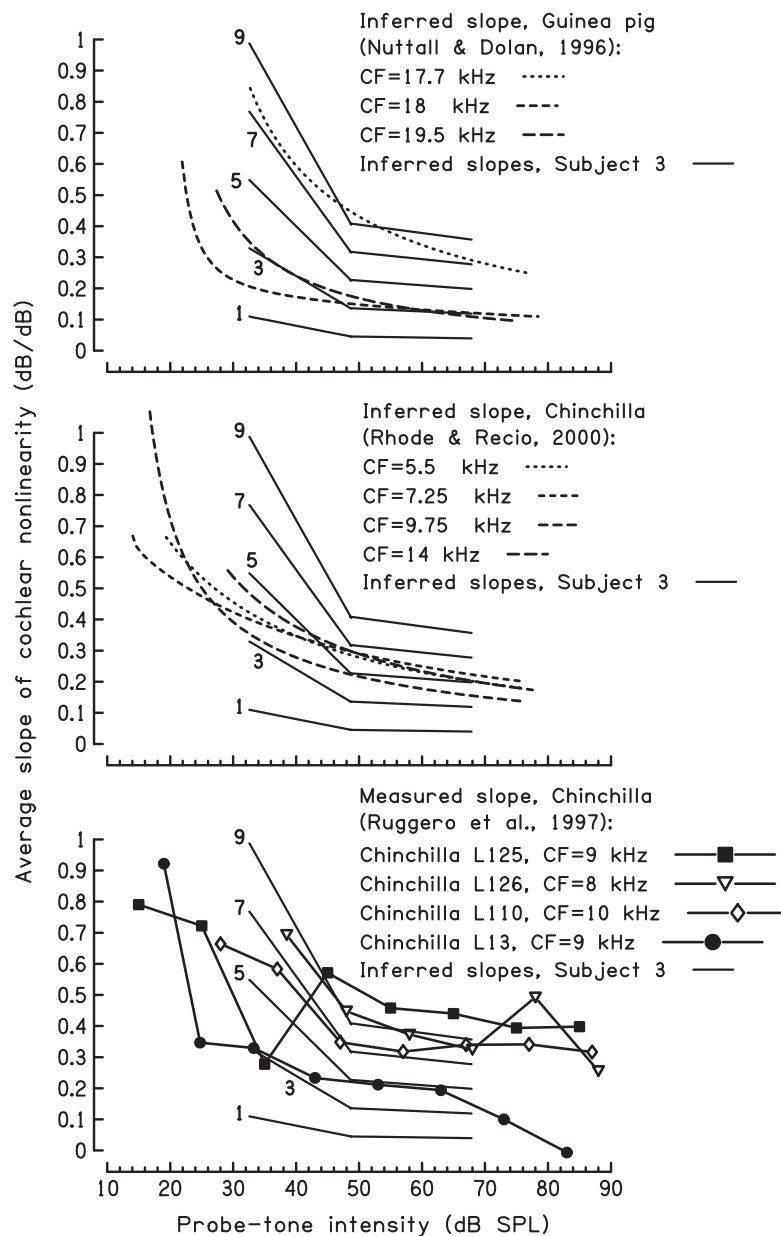


Figure 21 Comparison of inferred slopes of the cochlear nonlinearity: Subject 3 versus animals. The inferred slopes of the cochlear nonlinearity for Subject 3 are overlaid upon broken curves (Eq. 12 for upper panel, Eq. 14 for middle panel) or data (for lower panel) for the nonlinearity slopes from furred mammals. The bisegmented lines here are the same in the upper, middle, and lower panels, and represent different values of the unknown proportionality constant K for Subject 3. Rightward shifts, in decibels, were applied to the physiological curves or plots, as described in the caption to Figure 17.

25-ms tone, which was 20 times longer than the equivalent rectangular duration of 1.25 ms of the present probe-tone. Therefore, the plots based upon physiological data might, to a first approximation, all be moved rightwards by $10|\log_{10}(1/20)| = 13.01$ dB.

A different adjustment must be considered, however, because the ratio of total energies of two pure tones of identical amplitude within a given time interval is the ratio of their sine-wave frequencies^f. An upper limit of the needed shift of the animal-

based plots in Figures 17, 18, 19, 20 and 21 can hence be obtained by including the ratio of the two sine-wave frequencies as a multiplier. For example, for comparing results obtained with a 9 kHz tone in animals to the present results obtained with a 2 kHz tone, the actual needed shift might be as much as $10|\log_{10}[(1/20)(2/9)]| = 19.54$ dB.

However, such analyses of the present probe-tone's relative ability to drive the basilar membrane are unsophisticated; for example, they ignore the fact that the basilar membrane's stiffness changes with location (and hence with CF), such that its responsiveness to stimulation may differ by pure-tone frequency. Therefore, the decibel shift needed to compensate for differences in tone energy, shape, and duration between human and animal studies could conceivably be as little as that required to compensate for frequency differences, namely, $10|\log_{10}(2/9)| = 6.53$ dB. The range between the upper and lower respective estimates here of 19.54 dB and 6.53 dB is 13.01 dB, as it should be, calculations-wise. Hence, although we cannot prescribe the exact hypothetical shifts, we can at least bracket the values that should be needed in order to make the present psychophysically-derived cochlear nonlinearity slopes align with the physiologically-derived ones. The hypothetical allowed ranges are specified in the caption to Figure 17.

Ultimately, it was decided to shift the physiologically-derived nonlinearity slopes "by eye" until they best aligned with the psychophysically-derived curves for Subject 1A (Experiment 2). The caption to Figure 17 lists the shifts actually made. They are maintained throughout the remaining relevant figures, Figures 18, 19, 20 and 21. They prove to be well within the allowed limits, with one exception. That was the curve for CF=7.25 kHz (data of [15]), whose shift of 19.54 dB exceeds the maximum allowed shift of 18.60 dB for that frequency. This case is relatively unimportant, however, because a specific shift was hard to judge, due to the shallowness of the curve, and could easily have been set within the prescribed range.

Analysis (3): possible values of the unknown parameter

Figures 17, 18, 19, 20 and 21 suggest that, for nonlinearity slopes in humans to resemble those of animals, the unknown parameter K must be set between 3 and 20. Specifically, acceptable ranges of K are 5–20 for Experiment 2 (compare Figures 17 & 18 to Figure 16) and 3–7 for Experiment 1 (compare Figures 19, 20 and 21 to Figure 16). For Subjects 1A (Figure 17) and 2A (Figure 18), a few of the curves of psychophysically-derived nonlinearity slopes conform well to a few of the curves of physiologically-derived nonlinearity slopes. But for Subjects 1, 2, and 3 (Figures 19, 20 and 21, respectively), the steep upper line segment of the two-segment model of the nonlinearity *slope* conforms poorly to the curves for the physiologically-derived slopes. Indeed, the latter might have to be shifted horizontally to their maximum allowed limits in order for something resembling "overlap" to occur. The needed shifts are unclear, and so (as noted above) the shifts inferred from Figure 17 were employed, as "conservative" estimates. Note well that within the Schairer et al. [4] model, the slopes of the cochlear nonlinearity would simply be two discrete numbers, appearing in Figures 17, 18, 19, 20 and 21 as two disjoint horizontal lines.

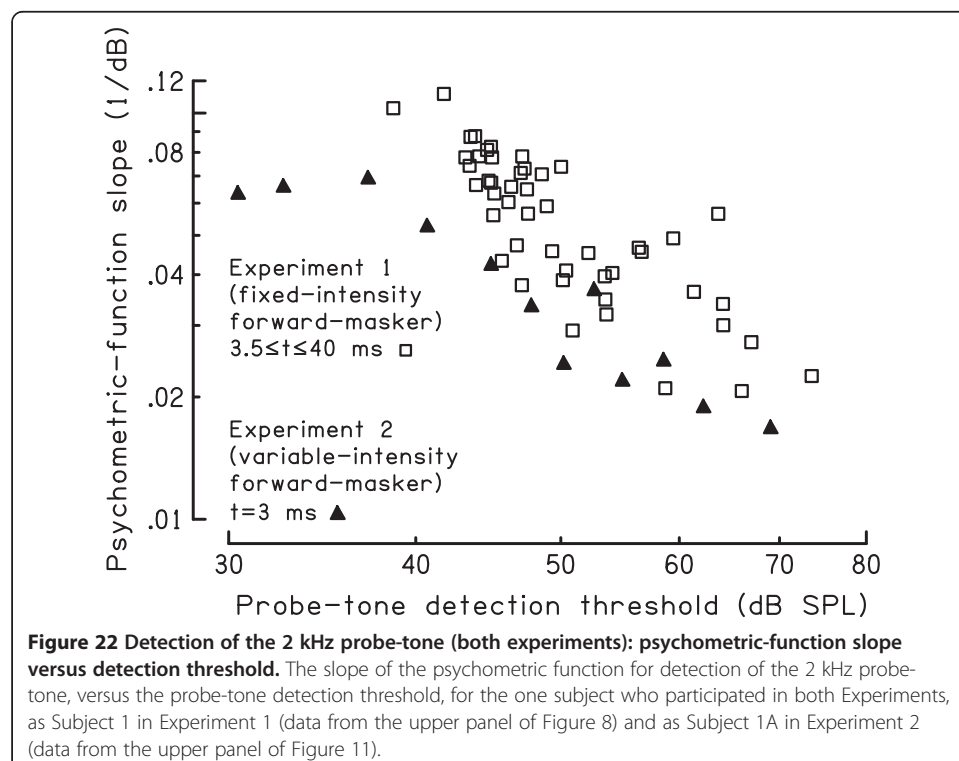
The slopes-of-nonlinearities predicted from the psychometric functions of Experiment 1 clearly differ from those predicted from the psychometric functions of Experiment 2. Which are correct?

Recall that one experimental subject had participated in both Experiment 1 (as Subject 1) and Experiment 2 (as Subject 1A). This offers an opportunity to compare the psychometric-function slopes from two different experimental methods of elevating the probe-tone's detection threshold. Figure 22 shows the psychometric-function slopes for Subject 1/1A. The symbols for Experiment 2 fall on the lower edge of the symbols for Experiment 1, implying that this individual performed differently in the two experiments. The laboratory, and the laboratory procedures did not change notably from Experiment 1 to Experiment 2; an explanation, therefore, must focus on the differences between the actual tasks. The most obvious speculation to make is that the closeness of the probe-tone to the termination of the forward-masker in Experiment 2 may have confounded the probe-tone's detection, the illustrated lower psychometric-function slopes representing wider psychometric functions and, within the Schairer et al. [4] model, a wider output distribution.

Analysis (4): the cochlear mechanical nonlinearity in man, inferred from its rate-of-change

Equations for the nonlinearity

Equations 6-10, which express the inferred average rate-of-change of the human cochlear nonlinearity with intensity x , can be used to infer the cochlear nonlinearity itself. They need only be integrated. Such integration must have a lower limit, here denoted x_{\min} in dB SPL, which logically must be the stimulus' absolute detection threshold ("quiet" threshold), i.e., its threshold in the absence of any kind of masking.



To quantify the cochlear mechanical nonlinearity in man, let us define a function $S(x)$ within the context of integrating the general form for $\bar{f}(x)$, Eq. 4:

$$F(x) = \int_{x_{\min}}^x \bar{f}(x) dx = K \cdot \int_{x_{\min}}^x s(x) dx = K \cdot [S(x) - S(x_{\min})] \quad dB \quad (15)$$

In Figure 8, psychometric-function slopes are described by a two-part function. Let the transition point between those parts be at an intensity x_{tran} . Then $\bar{f}(x)$ has two parts, for which

$$F(x) = \int_{x_{\min}}^x \bar{f}(x) dx = K \cdot \int_{x_{\min}}^x s_1(x) dx = K \cdot (S_1(x) - S_1(x_{\min})) \quad dB \quad (16a)$$

for $x \leq x_{\text{tran}}$ dB SPL, and

$$\begin{aligned} F(x) &= \int_{x_{\min}}^x \bar{f}(x) dx = K \cdot \left[\int_{x_{\min}}^{x_{\text{tran}}} s_1(x) dx + \int_{x_{\text{tran}}}^x s_2(x) dx \right] \\ &= K \cdot (S_1(x_{\text{tran}}) - S_1(x_{\min}) + S_2(x) - S_2(x_{\text{tran}})) \quad dB \end{aligned} \quad (16b)$$

for $x > x_{\text{tran}}$ dB SPL, where generally $S_1(x_{\text{tran}}) \neq S_2(x_{\text{tran}})$.

If we extend Eqs. 6a, 7a, 8a, 9, and 10 below their stated lower limits, right down to the absolute probe-tone detection threshold, and if we also solve for $S(x)$ or $S_1(x)$ or $S_2(x)$, then for Subject 1 we obtain

$$S_1(x) = -0.0085 \frac{x^2}{2} + 0.455x \quad dB, \quad 27.8 \leq x \leq 48.9 \quad dB \quad SPL \quad (17a)$$

$$S_2(x) = -0.0003 \frac{x^2}{2} + 0.054x \quad dB, \quad 48.9 \leq x \leq 73.6 \quad dB \quad SPL \quad (17b)$$

and for Subject 2 we obtain

$$S_1(x) = -0.012 \frac{x^2}{2} + 0.644x \quad dB, \quad 25.3 \leq x \leq 49.9 \quad dB \quad SPL \quad (18a)$$

$$S_2(x) = -0.0003 \frac{x^2}{2} + 0.06x \quad dB, \quad 49.9 \leq x \leq 74.0 \quad dB \quad SPL \quad (18b)$$

and for Subject 3 we obtain

$$S_1(x) = -0.004 \frac{x^2}{2} + 0.24x \quad dB, \quad 22.6 \leq x \leq 48.6 \quad dB \quad SPL \quad (19a)$$

$$S_2(x) = -0.0003 \frac{x^2}{2} + 0.06x \quad dB, \quad 48.6 \leq x \leq 67.7 \quad dB \quad SPL. \quad (19b)$$

These equations must be substituted, where appropriate, into Eqs. 16a and/or 16b. In contrast, following similar steps for Subject 1A yields

$$S(x) = -53 \frac{x^{-0.9}}{0.9} \quad dB, \quad 27.8 \leq x \leq 69.0 \quad dB \quad SPL \quad (20)$$

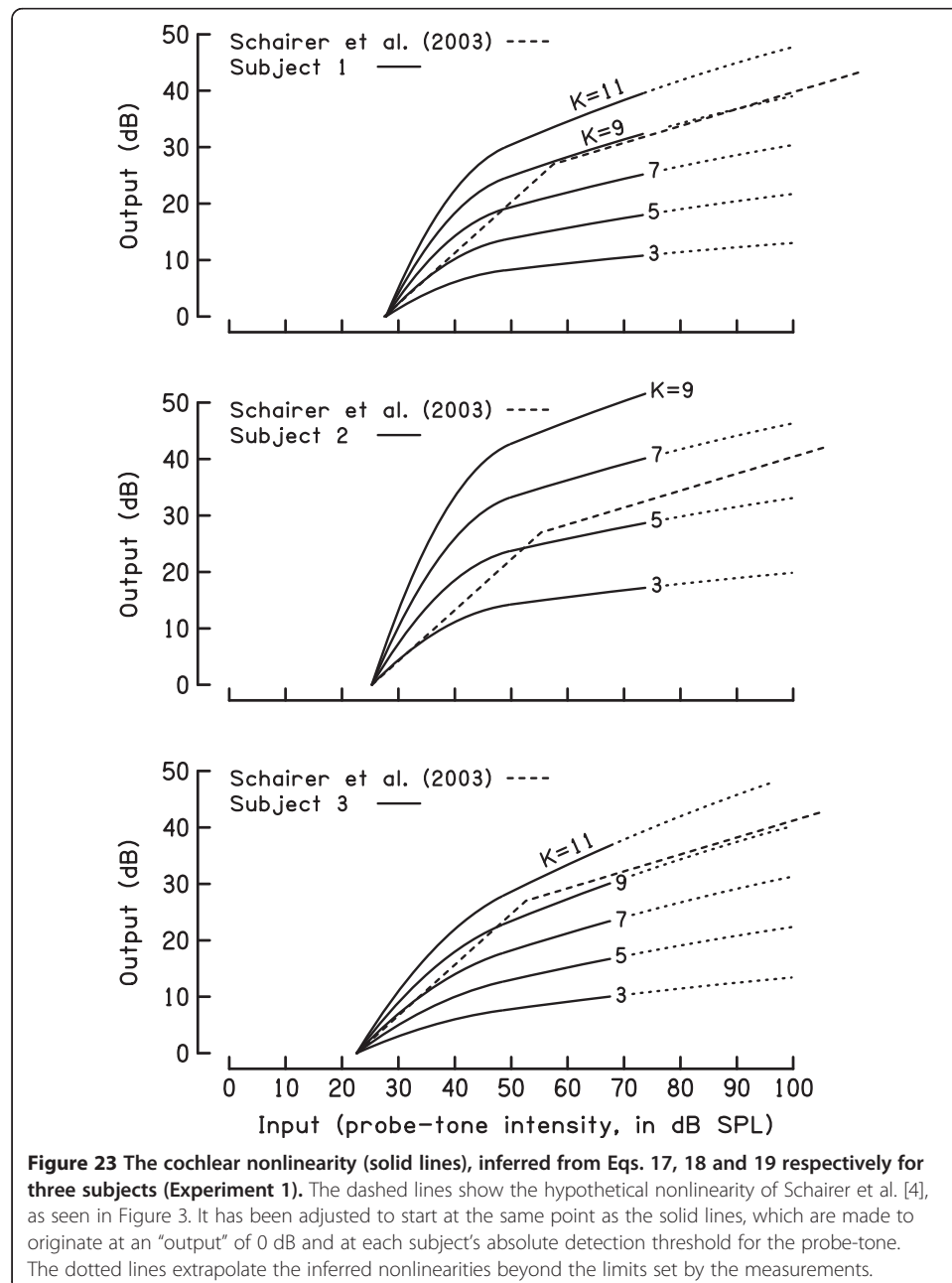
and for Subject 2A,

$$S(x) = -2.4 \frac{x^{-0.12}}{0.12} \text{ dB}, \quad 26.7 \leq x \leq 68.9 \text{ dB SPL}. \quad (21)$$

These equations are substituted into Eq. 15, producing *non-negative* values of the nonlinearity, as required.

The inferred cochlear nonlinearity in man

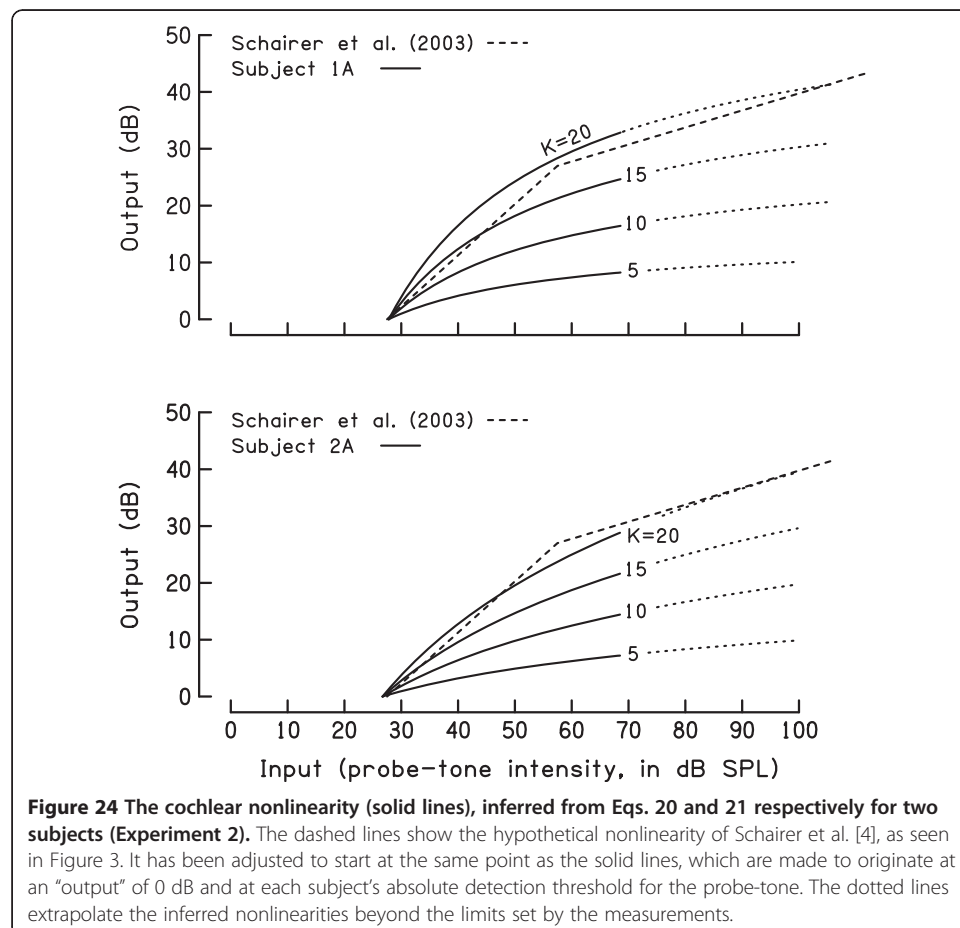
Figure 23 shows the cochlear nonlinearity generated by Eqs. 17, 18 and 19, for values of K that make the average slopes of the nonlinearity for humans comparable to those



for animals. The cochlear nonlinearity for the Schairer et al. [4] model is also shown, and has been adjusted to start at the same point as the curves.

Figure 23 demonstrates that the nonlinearities predicted from Experiment 1 have a continuum of similar shapes across subjects. Slopes at low intensities can be greater or lesser than the Schairer et al. [4] predicted slope of 1. A slope exceeding 1 represents profound amplification. Nonetheless, each of the curves for Subjects 1 and 2 (less so for Subject 3) bends smoothly at roughly 15–20 dB above its starting point, in contrast to the abrupt bend at 30 dB above the starting point which Schairer et al. [4] had proposed. Ironically, the regime of low thresholds represents weak forward-masking, such as that at the extreme right-hand-side of the postmasker recovery curve (i.e., Figure 4), or the extreme left-hand-side of the growth-of-forward-masking curve (Figure 9). Detailed psychometric functions for probe-tone detection at long postmasker recovery times are not known, nor are they known for forward-maskers that result in probe-tone detection thresholds very close to the probe-tone's absolute detection threshold.

The present model cannot accurately specify the point of bend or the sharpness of the bend, due to the computation being only approximate, as mentioned above. Indeed, a *sharp* bend may be illusory; the recorded nonlinearities for animals do not show a distinct point of change of slope. And that initial low-intensity rise has a slope below 1. Figure 24 shows the hypothetical cochlear nonlinearities inferred using Eqs. 20 and 21



(Experiment 2). The upper slopes of the inferred nonlinearities still resemble that of the Schairer et al. [4] model, but there are no distinct bends, unlike Figure 23.

Importantly, the ranges in decibels of output of the inferred nonlinearities in Figures 23 and 24 are of the same order of magnitude as those seen for animals, which are roughly 13–30 dB across animals [15–17].

Figure 25 shows the inferred cochlear nonlinearities for the one subject common to both Experiment 1 (“Subject 1”) and Experiment 2 (“Subject 1A”). Those different experiments result in clear differences in the inferred nonlinearity, because recovery from forward-masking (Experiment 1) provided different psychometric functions for this subject than obtained from growth of forward-masking (Experiment 2).

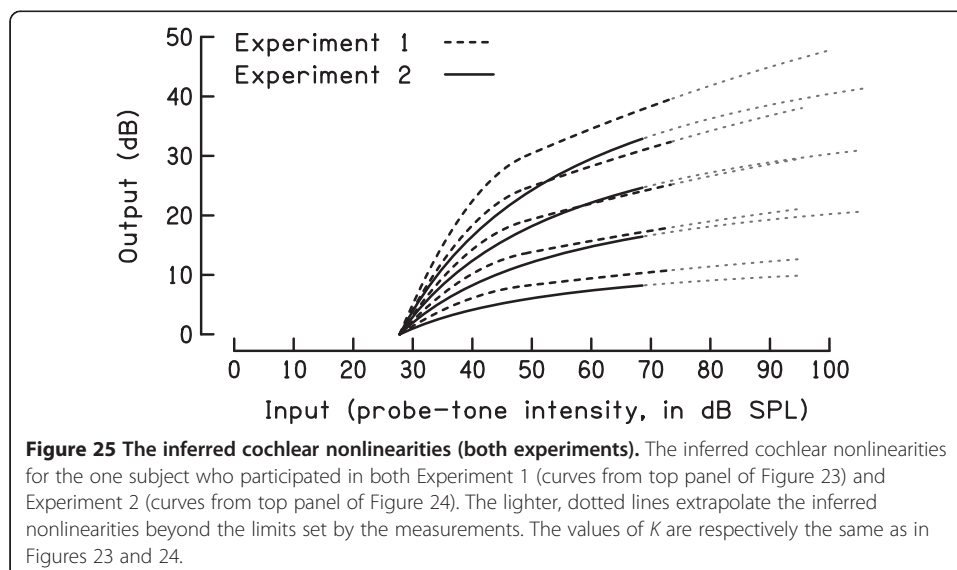
Discussion: the quality of the present predictions, and of the psychometric functions that they are based upon

The quality of the present predictions: how rigorous is the bisegmented nonlinearity in the Schairer et al. [4] model?

Figures 23 and 24 show hypothetical cochlear mechanical nonlinearities whose lower and upper branches do not precisely follow the Schairer et al. [4] model of the nonlinearity. In particular, the slopes of the lower branches of the nonlinearity in Figure 23 can exceed those of the Schairer et al. [4] model, whereas the slopes of the upper branches in Figure 23 can be lower than, higher than, or equal to those of the Schairer et al. [4] model. Similar but less extreme departures are seen in, or can be inferred from, Figure 24. Do such discrepancies invalidate the present nonlinearity computations based upon the results of Experiments 1 and 2? Hardly, as will now be shown.

The origin of the bisegmented nonlinearity: Yates et al. [7]

Schairer et al. [4] cite Yates et al. [7] as the source for the bisegmented nonlinearity used by Plack and Oxenham [6], although the latter do not explicitly mention Yates et al. [7] in their explanation of their model [6, p. 1599]. Plack and Oxenham [6] actually credit an earlier Oxenham paper for their model, which in turn cites a related paper by Yates.



Regardless, if Yates et al. [7] is indeed the source of the Plack and Oxenham [6] bisegmented cochlear nonlinearity, then some critical commentary is overdue, and proceeds as follows.

Yates et al. [7] illustrated some smoothed empirical firing-rate-versus-intensity plots (here called “rate-intensity functions”) for four neurons exposed to pure tones of frequencies at or below each respective neuron’s CF. For tones at CF, whose intensity is specified in dB SPL, Yates et al. [7] found rate-intensity functions which were either sigmoidal in shape, or “sloping-saturating”, or straight. “Sloping-saturating” means that the rate-intensity function shows a bend, then climbs for an additional 30 dB or more but with a notably lesser slope (see [20] for sources of some examples). Yates et al. [7] noted that rate-intensity functions for frequencies well-off-CF tend to be sigmoidal, with a central straight section, i.e., one that is linear in dB SPL. They noted also that basilar-membrane displacement is, likewise, empirically linear in dB SPL at tone frequencies well below a given spot’s CF. Yates et al. [7] hence assumed that the well-off-CF rate-intensity function represents the dependence of a neuron’s firing rate upon basilar-membrane motion. They then chose a single neuron whose on-CF firing was sloping-saturating, and whose well-off-CF firing was sigmoidal with a linear central section. Then, for each firing rate represented by a point on the linear central section of the well-off-CF rate-intensity function, Yates et al. [7] found the intensity giving the same firing rate on the respective on-CF sloping-saturating rate-intensity function, and assumed that the latter intensity in dB SPL corresponded to the dB of basilar-membrane motion presumed to drive the well-off-CF rate-intensity function. In this manner, Yates et al. assembled a derived basilar-membrane input/output function for each of their four aforementioned neurons. In their own words, “all [of these four derived input/output] curves take on some aspect of a general form: an initial slope of unity, indicating a linear relationship between SPL and BM [basilar-membrane] amplitude, turning over more-or-less abruptly to assume a second, straight, section with a slope of about 0.2-0.25” [7, p. 211]. This, presumably, is the origin of the 0.2 dB/dB slope adopted by Plack and Oxenham [6] and later adopted in turn by Schairer et al. [4]. Yates et al. [7] used several other neurons to provide similarly-derived examples of the inferred basilar-membrane input/output function.

Closer inspection of the Yates et al. [7] nonlinearity

There are, however, a few major problems with the Yates et al. [7] construction of the cochlear mechanical nonlinearity. These are important, because they excuse the differences evident in Figures 23 and 24 between the cochlear nonlinearity as presently computed and the model nonlinearity of Schairer et al. [4]. Regarding Yates et al. [7], the well-off-CF functions shown by Yates et al. [7] tend to have linear sections only 15 dB wide at most. Indeed, in their explanatory example, the linear section is merely 10 dB wide – just enough to span the “hinge” in their inferred cochlear nonlinearity, which in their illustration nonetheless has an *output* range of 30 dB, something of a liberty.

Also, the Yates et al. [7] method of constructing the cochlear nonlinearity inherently assumes that the bend in the rate-intensity functions of sloping-saturating neurons represents the bend in the nonlinearity itself. Yates et al. [7] do not conceal the source of this notion, namely, an overcited paper by Sachs and Abbas [21], who proposed an

approach to quantifying the effect of the nonlinearity on rate-intensity functions, an approach from which Yates et al. [7] clearly borrowed a great deal. For example, both papers inherently assume that basilar-membrane mechanical properties at one point along its length are mimicked at some other point, although in fact the mechanical properties change with location. Sachs and Abbas [21] also assumed a cochlear nonlinearity that had a slope of unity up to 73 dB SPL, above which the slope was 0.37, based upon others' early measurements of the cochlear nonlinearity in monkeys. That intensity of 73 dB SPL is now known to be far too high for the bend, but Sachs and Abbas (using cats) nonetheless successfully found sloping-saturating rate-intensity functions whose bends were at 73 dB SPL. Sachs and Abbas [21] also showed plots that suggest that the bend point in the rate-intensity functions for sloping-saturating neurons varies over 40 decibels across neurons! Elsewhere, Palmer and Evans [22] also noted a broad range for the bend, one of 20 decibels. Altogether, such numbers suggest that sloping-saturating rate-intensity functions are, as Palmer and Evans [22] noted, not a result of cochlear nonlinearity^g.

Altogether, the cochlear nonlinearity inferred by Yates et al. [7] must be considered a convenient contrivance. But Schairer et al. [4] and Schairer et al. [5], after Plack and Oxenham [6], had adopted the Yates et al. [7] nonlinearity. As such, deviations from it in Figures 23 and 24 should not be regarded as a failure of the present computations, but rather, of the underspecification of the nonlinearity in [4-7].

The quality of the present psychometric functions: (1) the advantages of the present method of obtaining psychometric functions over that of Dai [23]

Schairer et al. [4] and by Schairer et al. [5] used a method of Dai [23] to obtain psychometric functions through adaptive tracking, and to find their slopes. Dai [23, p. 3135] had concluded that adaptive tracking is "a better choice than the constant-stimulus method for measuring psychometric functions". However, Dai's approach is actually the less desirable one, as follows.

Does adaptive-tracking really yield psychometric functions which are more precise than those obtained through the method of constant stimuli?

Dai's [23] conclusion was largely based upon computer simulations which he did to imitate a hypothetical observer performing anywhere from 120 to 900 trials, in 60-trial blocks, with each successive block starting at the threshold that had been identified in the previous block. (The use of small blocks brings its own problems; see below.) Dai's [23] approach would seem to allow listener-experience-based improvement in the estimated threshold, but it contrasts to actual experiments, in which starting conditions might be the same for each block. Regardless, Dai [23] simulated three experimental methods: (1) adaptive tracking using the 2-down 1-up rule or (2) the 3-down 1-up rule; and (3) the method of constant stimuli. Step size was a parameter of the simulations. The "true psychometric function of the simulated observer, which was used to generate the responses" [23, p. 3136] was a cumulative Gaussian in the variable $d'/\sqrt{2}$, where d' is the Signal Detection Theory index of detectability [24]. Dai [23] defined d' in terms of "signal level" x as $d' = (x/\alpha)^\beta$ for parameters α and β , which were to be estimated after-the-fact from the simulations, in which the actual chosen values of α and β were $\alpha = 1$ and $\beta = 1$. Note well that Dai's x has intensity units, not decibel units. Dai's psychometric function *for data*

generation was a cumulative Gaussian in the intensity $x/\sqrt{2}$. In contrast, the present psychometric functions are cumulative Gaussians in x in dB SPL, a significant difference, as will be explained below. Nonetheless, both Dai [23] and the present work fitted functions to percentages-correct by minimizing the same weighted sum-of-squares-of-residuals, i.e., that used in Probit Analysis [1].

Dai [23] found that the generated estimates of α and β converged towards the true values as the number of trials approached 900, and as the step size increased from 1 dB to 12 dB. But the predicted β proved especially sensitive to step size, as had been found elsewhere (citations in [23]); in particular, for a step size of 1 dB, and using just 120 trials, the β obtained under adaptive tracking diverged significantly from its true value. The β computationally obtained by Dai [23] under the method of constant stimuli, however, did not. Dai's [23] recommended solution to the divergence was to encourage the use of larger step sizes.

However, in earlier work from the Jesteadt laboratory [19,25], the present author had tried blocks of adaptive tracks using step sizes of 8 dB followed by step sizes of 4 dB, and had found that roughly 1 out of every 4 adaptive tracks had to be discarded because it produced unrealistically low detection or discrimination thresholds. That is, subjects were able to make lucky guesses, which, due to the step sizes, took their thresholds down to absurdly low stimulus levels from which 50-trial blocks did not allow enough trials to recover. Of course, an experimenter might try to overcome this problem by using many more trials in each single block; in that case, one might just as well use the method of constant stimuli! Regardless, the point is that simulated psychophysical performances and *actual* psychophysical performances can give startlingly different outcomes.

Does adaptive tracking account for learning?

Dai himself [23, p. 3135] raised an important point when he noted that "The ability to trace the underlying psychometric function is particularly desirable when the performance of the observer undergoes a marked change as a result of learning, fatigue, fluctuation of attention, etc.". *Learning* is often presumed to reflect reduction of internal noise. Unfortunately, the adaptive-tracking method itself may not take learning into account well, as no learning was mentioned in Schairer et al. [4] or in Schairer et al. [5], and no learning by listeners was evident in other papers from the same laboratory, papers concerning detection or discrimination (e.g., [13,19,25,26]). In contrast, substantial learning was evident in the experiments reported here, manifested as shifts of the psychometric function to lower and lower SPLs over successive days of testing; a single complete psychometric function for a given time gap was produced by each subject on each testing day. In the present Experiments 1 and 2, learning effects were paramount despite subjects' differences in the degree of previous laboratory listening experience (i.e., none for Subject 1/1A and Subject 2A, some for Subject 3, and very much for Subject 2). In each new stimulus condition of Experiments 1 and 2, improvement of detection threshold of as much as 10 dB occurred from the first to the second day's trials, with successively lesser improvements over the following one or two days, just as found elsewhere for detection of longer tones [27,28]. In Experiment 1, occasional within-day retesting with multiple probe-tone intensities showed that daily detection performance had asymptoted, but nonetheless, over successive days it continued to improve (see also [27-29]). As time gap in Experiment 1 changed from week to week, there was an imperceptible decline in the *range* of improvement in percentages-

correct over successive days *for a given time gap*, an improvement seen elsewhere with another constant-intensities two-alternative forced-choice task [30] involving comparable time (months) and practice (thousands of trials). Of course, learning still re-occurred at each new time gap, just as found in experiments where a detected tonal frequency was changed [27]. Altogether, the learning effects noted here reflect those noted over a variety of much earlier studies, not all of which employed two-alternative forced-choice. Evidently, then, if “the underlying psychometric function is particularly desirable when the performance of the observer undergoes a marked change as a result of learning” [23, p. 3135], then such a psychometric function cannot be obtained through adaptive tracking, which cannot therefore be “a better choice than the constant-stimulus method for measuring psychometric functions” [23, p. 3135].

The advantage of accounting for learning is evident in the actual forward-masked detection thresholds in Experiment 1. The latter, when averaged across the three subjects there (graphed in [9]), are as much as 10 dB lower than thresholds for recovery from forward-masking obtained elsewhere with comparable stimuli ([31], two-alternative forced-choice tracking; [32], Bekesy tracking), despite the fact that the average age of the three subjects in Experiment 1 was 32, higher than the twenty-something average age which is typical of student-listener cohorts and hence likely to result in higher thresholds. The 10-dB difference is comparable to the circa-10-dB threshold drop seen in Experiments 1 and 2 due to learning, and cements the notion that tracking methods may not account well for learning.

Does Dai's method give a better estimate of the slope of the psychometric function?

Last but hardly least, there is the practical issue of the actual fit of the psychometric functions to the percentages-correct. The psychometric functions fitted by Schairer et al. [4] and by Schairer et al. [5] were cumulative Gaussians in $d' = (x/\alpha)^\beta$, after Dai [23], where x has intensity units (rather than decibels). The discrete values of x used in fitting functions to percentages-correct depends upon the technique that is used to establish the probe-tone's detection threshold. Adaptive tracking, like the method of constant stimuli, focuses on percentages-correct (through correct/incorrect criteria), determining stimulus intensities indirectly. But the experimenter using the method of constant stimuli can specify the stimulus intensities employed and how often they are used; however, in an adaptive track, the situation is not so flexible, because the set of stimuli used and their frequencies of occurrence depend upon (1) the tracking method itself, and (2) the actual performance of the subject. The percentages-correct in Dai [23] and in Schairer et al. [4] and in Schairer et al. [5] can be examined by drawing a horizontal line at the 75%-correct mark in each graph and counting the number of data points which are above or below that line. For the three experiments of Schairer et al. [4] and the three experiments of Schairer et al. [5], all illustrations but those for the second experiment of Schairer et al. [4] (“variable-signal” with maskers of 60 dB SPL or of 90 dB SPL) reveal the data plots to be top-heavy. That is, the number of discrete values of the intensity which produce percentages-correct above 75% exceeds the number of discrete intensity values which produce percentages-correct below 75%. This is especially evident in Schairer et al. [5] thanks, ironically, to a two-track adaptive procedure intended to “obtain a larger range of PCs [percentages-correct] for the PF [psychometric function] fits” [5, p. 2199], which provided a greater number of employed intensities. With more data points above the 75% line

than below it, the fit of the psychometric function will, *regardless of the employed weighting scheme*, focus on the uppermost data points, thus tending to underestimate the psychometric-function slope *at any percentage-correct along the curve*. Further, the upper data points may be more heavily weighted in the quantity that was actually minimized in the curvefitting. That quantity was a sum [23, Eq. 3] in which each term contains a multiplicative weight which is the number of trials associated with a particular percentage-correct. That number of trials may have been larger for the higher percentages-correct if the associated stimulus intensities had been visited more frequently in the adaptive track (s) than the lower stimulus intensities.

In contrast, the graphs of percentages-correct for the present Experiments 1 and 2 are not top-heavy. Hence, when the psychometric data and fitted functions for the present Experiments 1 and 2 were plotted versus intensity in dB SPL, slope at 75% was not underestimated, thanks to a close fit of the *symmetric* function to the data in coordinates of percentage-correct versus dB SPL. See Figure 10, whose fits of function to data are imitated by similarly good fits for Experiment 1, which the diligent reader can find in [8,9,14]. In those same plotting coordinates, however, the psychometric functions of Schairer et al. [4] and of Schairer et al. [5] are asymmetrical, with greater acceleration in the upper halves, as evident through close inspection of Figures 3 and 10 of Schairer et al. [4] and Figures 3 and 6 of Schairer et al. [5]. Such asymmetry de-emphasizes the fitting of the psychometric function to the lower data points, contributing to the misestimation of the slope of that function at any percentage-correct, and clearly underestimating the slope at 75% and below. The asymmetry is presumably due to using $d'/\sqrt{2}$ (after [23]) as the independent variable in the psychometric function, rather than using the stimulus intensity in dB SPL. However, it is dB SPL, not $d'/\sqrt{2}$, which is the intensity measure of interest in the Schairer et al. [4] model (Figure 3).

Dai [23] had, besides running simulations, obtained *empirical* just-noticeable *frequency* differences using adaptive tracking with either a 2-down 1-up rule or a 3-down 1-up rule or the method of constant stimuli. In each case, Dai's [23] fitted psychometric functions more closely followed the upper portions of the percentages-correct, such that the psychometric-function slope at 75% correct was underestimated – the same problem that is evident in Schairer et al. [4] and in Schairer et al. [5]. His fitted curves, too, are similarly asymmetric^h.

The quality of the present psychometric functions: (2) why they offer unprecedented precision

The simulations of Garcia-Perez [33]

An important aspect of the present forward-masked detection thresholds is that the conservative and painstaking methods employed to obtain them rendered them of unprecedented precision. Garcia-Perez [33] confirms the unprecedented precision of the present results – by showing the *lack* of precision of detection thresholds obtained by the favored method used for obtaining detection and discrimination thresholds, namely, adaptive tracking. The latter is popular because of its greater speed than the method of constant stimuli, and it was the method employed by Schairer et al. [4] and by Schairer et al. [5], among others. Adaptive tracking typically employs equal dB steps when intensities are

adjusted up or down during the adaptive track. Tracking can follow a number of different rules; the rule of dropping the intensity after two correct identifications of the target stimulus and raising it after one incorrect identification, called 2-down 1-up, has been especially popular and has been used in countless papers, including many from the Jesteadt laboratory (e.g., [4,5,25,26]). Garcia-Perez [33] studied the relative efficacy of different tracking rules in m -alternative up-down tracking, where in the citations just mentioned, $m=2$, perhaps the most mundane choice. What Garcia-Perez [33] did was to generate simulated percentages-correct for detection or discrimination tasks, using either Weibull or logistic equations as the source psychometric functions. Those particular sigmoidal functions are justified from many psychophysical studies; recall that the cumulative Gaussians used in Probit Analysis are themselves approximations, as no exact sigmoidal solution exists. Garcia-Perez's findings for detection are the relevant ones here, and as such his findings for discrimination, though similar, will be ignored.

In performing his simulations, Garcia-Perez [33] incorporated a factor crucial to the present paper, namely, "the spread σ [*sic*] of a psychometric function". His definition of σ , however, differs from the present one (Eq. 3); indeed, his σ was described as "the width of the range of stimulus levels where ψ [the psychometric function] shows non-asymptotic behavior" [33, p. 2100]. The latter width was defined by Garcia-Perez using a mathematical rule involving a parameter chosen to give a σ that was "the width of the central 98% span of ψ " [33, p. 2100]. Garcia-Perez ran simulations of two-alternative forced-choice staircases, each of which was run until 200 reversals had occurred; from these the average of the last 180 reversals was taken as the detection threshold. Garcia-Perez used those to examine the behavior of the "landing point", that is, "the percentage-correct point on which the staircase converges" [33, p. 2101] under any particular adjustment rule and final step size and ratio of final step size to σ . That is, for each combination of conditions, he ran 5,000 replications, in order to obtain mean values and standard deviations of the detection thresholds (although he did not discuss their actual distributions). Each mean threshold was substituted back into the generating psychometric function to get the landing point; the standard deviations were likewise used to establish error bars on each landing point.

Results of the simulations of Garcia-Perez [33]

Garcia-Perez [33] noted that the *theoretical* landing-points stated in the literature, such as 70.7% for two-alternative forced-choice under the 2-down 1-up rule with equal up and down steps, all assume an infinite number of reversals, as well as infinitesimally small steps. In practice, as Garcia-Perez [33] discovered, the landing points tended to deviate downwards from their theoretical values, this difference tending to increase with increase in the ratio of final step size to σ , that is, as step size becomes a greater proportion of the width of the psychometric function. For example, rather than being 70.7%, the landing point for 2-down 1-up could be lower than 60% if the final step size was greater than, say, 0.35σ . Garcia-Perez [33] also realized that 200 reversals was more than typically used in psychophysical studies; he therefore repeated his simulations for the 2-down 1-up rule when threshold was determined from the last 10 of 12 reversals or from the last 40 of 42 reversals. The error bars associated with the landing points became even larger as the number of reversals decreased, and were of unequal size, being larger toward higher

landing points. Garcia-Perez's [33, p. 2104] overall conclusions bear repetition, and are best expressed in his own words:

The consequences of the differential bias of conventional up-down staircases may range from eliminating an actual difference in threshold to producing it when none was actually there, contingent on which up-down rule was used and how the spread of the psychometric function varies across conditions. The magnitude of this misestimation can only be determined if the spread of the psychometric function has also been estimated with sufficient accuracy, but this is rarely done in experiments designed to obtain quick threshold estimates via up-down staircases.

Implications for the precision of the present work relative to that of Schairer et al. [4] and Schairer et al. [5]

Schairer et al. [4] used 2-down 1-up adaptive tracking with final step sizes of 4 dB and blocks of 50 forced-choices, which altogether, according to Garcia-Perez [33], would introduce substantial variability into the landing point. That is, if the true landing point was 60% for a given forward-masking condition, rather than the theoretical 70.7%, then the true threshold would have been higher, perhaps by several decibels. The empirical narrowing of the psychometric function with either larger masker-probe time-gap or with lesser forward-masker intensity would (given a consistent final step size) systematically decrease σ , and therefore increase the ratio of final step size to σ , thus systematically increasing the divergence of the adaptive-tracking-derived threshold from its true measure. In the present Experiments 1 and 2, in contrast, the intensities used were no more than 2 dB apart (see Methods), which meant that the broader the psychometric function was, the greater was the number of different stimulus intensities used to establish the threshold. This resulted in confidence intervals for each threshold which were of roughly equal size across thresholds.

Altogether, the thresholds obtained in Experiments 1 and 2 should be far more precise than those obtained by Schairer et al. [4] and by Schairer et al. [5].

Why adaptive tracking may be even less precise than indicated by Garcia-Perez [33]

The thresholds obtained from the present Experiments 1 and 2 may be even more precise than those of Schairer et al. [4] and of Schairer et al. [5], for the following reason. Garcia-Perez [33] defined the width of a psychometric function for detection according to the kind of schemes which have been popular amongst auditory physiologists for defining the “dynamic range” of a primary afferent neuron. Those schemes defined dynamic range as the width of the sigmoidal rate-level function fitted to the plot of firing-rate-versus-intensity of the neuron [20]. But such schemes do not provide an *operational* measure, i.e., do not provide the *useful* stimulus-intensity-encoding range (in dB) of the neuron [34], which may be much narrower. By the same token, the criterion width used by Garcia-Perez [33] for psychometric functions, “the width of the central 98% span of ψ ”, is an extremely generous measure [20]; a more conservative measure can produce a much smaller “width” of a sigmoidal function, which would increase the ratio of final step size to σ , thereby *increasing* the deviation of any landing point from its theoretical value, hence *reducing* the precision of the thresholds inferred from landing points obtained using adaptive tracks.

"Fine structure" not seen in other studies

Finally, regarding Experiment 1, Figure 4 shows an unexpected rise in the probe-tone detection threshold circa $t=7$ milliseconds [8]. Studies of the recovery of the threshold of comparable stimuli from forward-masking [31,32,35] do not reveal this feature. Nonetheless, it is robust, being statistically significant as well as being associated with a sudden, momentary steepening of the psychometric function (see Figures 5, 6 and 7). It may be that learning, as well as greater precision, is required for this feature to appear.

Summary and conclusions

The present data confirm the general principles of the Schairer et al. [4] model of the influence of the cochlear nonlinearity upon psychometric functions for forward-masked probe-tone detection. Testing the Schairer et al. [4] model depends upon reliably documenting the slopes of the psychometric functions for forward-masked probe-tone detection, over a broad range of probe-tone detection thresholds. Such a range can be provided by strongly forward-masking a probe-tone, so that its detection threshold will be highly elevated at very short time-gaps between the constant forward-masker and the probe-tone, but will decline with increasing masker-probe time gap. That was Experiment 1, which provided psychometric functions for the detection of very brief Gaussian-shaped 2-kHz probe-tones. The dependence of psychometric-function slope upon probe-tone detection threshold could be approximated for each subject by a pair of conjoined non-horizontal line segments. For each subject, the point of junction was 21–25 dB above the probe-tone threshold in quiet. The nonlinearities predicted from such relations are increasingly compressive with increasing probe-tone intensity.

In Experiment 2, the masker-probe time-gap was fixed at 3 ms, just beyond the range of physical overlap of forward-masker and probe-tone. With increasing forward-masker intensity, the probe-tone detection threshold rises monotonically, as generally seen in the literature and as found by Schairer et al. [4, Figures 2 & 6] and later by Schairer et al. [5, Figure 2]. Also, the psychometric functions generally widen, as found by Schairer et al. [4] and later by Schairer et al. [5]. The decrease in psychometric-function slope with rising probe-tone detection threshold is adequately fitted by power functions. Hence, if psychometric-function slope is indeed determined as hypothesized by Schairer et al. [4], then the slope of the nonlinearity itself smoothly decelerates with increasing sound-pressure-level over roughly 20-80 dB SPL, showing no "elbow".

Power functions can also be applied to psychometric-function slopes versus probe-tone detection thresholds obtained for the growth of forward-masking by Schairer et al. [5]. Yet again, the cochlear nonlinearity is implied to be smoothly increasingly compressive with increasing probe-tone intensity, showing no "elbow". The dependence of psychometric-function slope upon probe-tone detection threshold was also examined for data gathered by Schairer et al. [4]. The cochlear nonlinearity is (once again) predicted to be smoothly increasingly compressive with increasing probe-tone intensity, showing no "elbow", as found from empirical measurement of basilar-membrane motion in small mammals.

Experiments 1 and 2, and re-analysis of the Schairer et al. [4] and the Schairer et al. [5] data, in whole and in parts substantiate the Schairer et al. [4] model. That model was subsequently extended here in order to reveal the human cochlear mechanical

nonlinearity itself. The average slope of the cochlear nonlinearity over some span of decibels centered on a particular intensity proves to be directly proportional to the slope of the psychometric function for forward-masked probe-tone detection which is centered at that intensity. Therefore, quantifying psychometric-function slope at its midpoint as a function of probe-tone detection threshold (i.e., the probe-tone intensity corresponding to the psychometric function's midpoint) leads to a further equation, for the average slope of the cochlear nonlinearity with intensity, in one unknown multiplicative parameter. The cochlear nonlinearity's actual slope can be obtained in animals; plotting theoretical cochlear-nonlinearity slopes for humans versus actual slopes for animals allows comparisons which suggest appropriate values of the unknown parameter.

The slopes of the psychometric functions themselves for probe-tone detection differ between Experiment 1 (recovery from forward-masking) and Experiment 2 (growth of forward-masking), hence the slopes-of-nonlinearity predicted from Experiment 1 in fact differ from those predicted from Experiment 2. Indeed, one subject participated in both experiments, and his psychometric-function slopes for Experiment 2 lie at the lower limit of those for Experiment 1. This begs the question of which of the respective experimental methods – varying the masker-probe time gap, or varying the forward-masker intensity – yields “correct” psychometric functions for inference of the cochlear nonlinearity.

The equations for the average slope of the cochlear nonlinearity can be integrated to give the nonlinearity. When the nonlinearity is predicted from Experiment 1, each computed nonlinearity bends smoothly at roughly 15–20 dB above its starting point. This differs from the Schairer et al. [4] model, which posits an abrupt bend, circa 30 dB above the starting point. The cochlear nonlinearities predicted from Experiment 2 show no distinct point of bending, instead resembling animal recordings. For both Experiments 1 and 2, the range (in decibels) of output of the inferred nonlinearity, from minimum to maximum, is similar to those seen in animals. In total, the Schairer et al. [4] model can be extended to provide credible inferences of the cochlear nonlinearity.

Endnotes

^a It is important to note that increasing compression with intensity, for low-to-moderate intensities, has been found for all stimulus durations employed in furred mammals, including clicks (brief impulses, typically 0.1 millisecond long). Therefore, all of the arguments made so far will be assumed to be independent of probe-tone duration.

^b “Precision” (in the sense of *consistency*) is used rather than “accuracy”, because judging accuracy requires comparison to “true” values of forward-masked thresholds, which are unknown. The proper use of, and interpretation of, such descriptive terms is an ongoing problem in biomedical work, one whose breadth is not to be underestimated [36].

^c What constitutes a “good fit” here is a geometrical exercise, not a computational one; the human eye is still the final arbiter of “fit”, as there is no *rigorously proven* general objective measure of fit.

^d The reader might confuse such velocities, as rates-of-change of displacement, with the rates-of-change of the cochlear input-output response. The former have units of

displacement per unit time, whereas the latter have units of [decibels of] displacement per [decibels of] intensity.

^e The latter points suggest an upturn in D for intensities above roughly 75 dB SPL. Such a higher-intensity upturn is noted in reviews [2,3]. It is presently unimportant, however, because Experiment 1 provides few detection thresholds in that regime, and Experiment 2 provides none at all. Such detection thresholds would require forward-masker intensities beyond what many experimental subjects (and Institutional Review Boards) might find tolerable.

^f Any pure tone has the same average power (i.e., energy/unit time) over a single period of its sine-wave, given a common amplitude [37, p. 26]. But the number of periods within a given time interval increases with increase in sine-wave frequency, such that tones of higher sine-wave frequency have inherently more average power.

^g There is further evidence that sloping-saturating rate-intensity functions are not a result of cochlear nonlinearity. Note first that sloping-saturating rate-intensity functions have been found not just in small furred mammals but also, in vastly different proportions, in birds, frogs, and even insects (the references are too numerous to mention), the latter not being considered to have the active amplification mechanisms associated with the cochlear nonlinearity [2,3]. Further yet, sloping-saturation is found for auditory neurons beyond the periphery (citations in [34]). Even within one well-studied species, the cat, there is uncertainty about what proportion of peripheral rate-intensity functions are sloping-saturating; estimates vary from 0% [38] to 9% [39] to 50% [21]. Given that sigmoidal rate-intensity functions are ubiquitous (again, the references are too numerous to mention; see citations in [20]), all of this suggests that sloping-saturation could be an experimental artifact. We must also ask whether the upper branch of the sloping-saturating rate-intensity function is useful at all in encoding changes in stimulus intensity. Palmer and Evans [40] found that for 18 sloping-saturating neurons in cats, the average slope of the upper-intensity limb was 1.31 spikes/second-dB, in contrast to 5.64 spikes/second-dB for the lower-intensity limb. Nizami [34] applied Signal Detection Theory to deduce the encoding ability in cats of 62 sigmoidal or sloping-saturating rate-intensity functions, and found that the latter had no advantage over the former. This suggests that sloping-saturating functions cannot reflect cochlear nonlinearity, otherwise good discriminability for intensity change would not be possible over the majority of the hearing range (if cats are any model for man).

^h There is another (and probably lesser) possible contributor to the better fit of psychometric functions to the present percentages-correct, namely, that the stimuli of Experiments 1 and 2 were given across blocks in descending order of intensity, such that the subject's frustration with the task, due to incorrect responses (feedback was continuously given), increased gradually and monotonically, rather than shifting back and forth in a manner that could evoke confusion, as during an adaptive track. Thus the subjects would be better prepared to detect the lowest-intensity stimuli employed – causing a more gradual drop in the percentages-correct with drop in probe-tone intensity, and hence a better fit in Probit Analysis.

Competing interests

The author declares no competing interests.

Authors' information

LN has a BSc in Physics, an MSc in Theoretical Physiology, and a PhD in Perceptual Psychology, all from University of Toronto. He is the author of 19 peer-reviewed papers, 13 Proceedings, 3 book chapters, and 11 letters-of-concern to editors, all in print format.

Acknowledgments

Writing self-funded. My special thanks to the experimental subjects for their enduring patience. Bruce A. Schneider (U. Toronto in Mississauga) sponsored the experiments through an NSERC grant. Doug Creelman (St. George Campus, U. Toronto) suggested the use of Probit Analysis. Walt Jesteadt (Boys Town National Research Hospital) suggested the re-examination of the forward-masked thresholds of Nizami [9] and kindly provided the comparison data of Schairer et al. [4] and of Schairer et al. [5]. Prof. Claire S. Barnes (VA Palo Alto HCS) provided valuable suggestions during proofreading. Professors Alfred L. Nuttall (Oregon Health and Science U.), William S. Rhode (U. Wisconsin), and Mario A. Ruggero (Northwestern U.) equally and kindly contributed their original records of the cochlear nonlinearity in animals. Finally, I thank the two anonymous reviewers for their very helpful critiques, and I thank Dr. Jiri Wackermann of the Institute for Frontier Areas of Psychology and Mental Health (Freiburg, Germany) for encouraging me to publish this paper.

Received: 23 January 2013 Accepted: 9 May 2013

Published: 9 May 2013

References

1. Finney DJ: *Probit Analysis*. London: Cambridge University Press; 1971.
2. Ulfendahl M: **Mechanical responses of the mammalian cochlea**. *Prog Neurobiol* 1997, **53**:331–380.
3. Robles L, Ruggero MA: **Mechanics of the mammalian cochlea**. *Physiol Rev* 2001, **81**:1305–1352.
4. Schairer KS, Nizami L, Reimer JF, Jesteadt W: **Effects of peripheral nonlinearity on psychometric functions for forward-masked tones**. *J Acoust Soc Am* 2003, **113**:1560–1573.
5. Schairer KS, Messersmith J, Jesteadt W: **Use of psychometric-function slopes for forward-masked tones to investigate cochlear nonlinearity**. *J Acoust Soc Am* 2008, **124**:2196–2215.
6. Plack CJ, Oxenham AJ: **Basilar-membrane nonlinearity and the growth of forward masking**. *J Acoust Soc Am* 1998, **103**:1598–1608.
7. Yates GK, Winter IM, Robertson D: **Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range**. *Hear Res* 1990, **45**:203–219.
8. Nizami L, Schneider BA: **The fine structure of the recovering auditory detection threshold**. *J Acoust Soc Am* 1999, **106**:1187–1190.
9. Nizami L: **On auditory dynamic range**. *PhD thesis*. University of Toronto, Psychology Department; 1999.
10. Nizami L: **The human cochlear mechanical nonlinearity inferred through the Schairer et al. (2003) model**. In *Fechner Day 2012: proceedings of the 28th annual meeting of the International Society for Psychophysics, October 2012*. Edited by Leth-Steensen C, Schoenherr JR. Ottawa, Ontario, Canada: International Society for Psychophysics; 2012:12–17.
11. Gabor D: **Theory of communication**. *J Inst Elec Eng London* 1946, **93**:429–457.
12. Schneider BA, Pichora-Fuller MK, Kowalchuk D, Lamb M: **Gap detection and the precedence effect in young and old adults**. *J Acoust Soc Am* 1994, **95**:980–991.
13. Nizami L: **Threshold vs. duration for Gaussian-shaped tone-pips of one to four periods duration**. *Percept Motor Skills* 2004, **99**:821–836.
14. Nizami L: **Afferent response parameters derived from postmasker probe- detection thresholds: 'the decay of sensation' revisited**. *Hear Res* 2003, **175**:14–35.
15. Rhode WS, Recio A: **Study of mechanical motions in the basal region of the chinchilla cochlea**. *J Acoust Soc Am* 2000, **107**:3317–3332.
16. Nuttall AL, Dolan DF: **Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig**. *J Acoust Soc Am* 1996, **99**:1556–1565.
17. Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L: **Basilar-membrane responses to tones at the base of the chinchilla cochlea**. *J Acoust Soc Am* 1997, **101**:2151–2163.
18. Dallos P, Popper AN, Fay RR: *The Cochlea*. New York: Springer-Verlag; 1996.
19. Nizami L, Reimer JF, Jesteadt W: **The intensity-difference limen for Gaussian-enveloped stimuli as a function of level: tones and broadband noise**. *J Acoust Soc Am* 2001, **110**:2505–2515.
20. Nizami L: **Estimating auditory neuronal dynamic range using a fitted function**. *Hear Res* 2002, **167**:13–27.
21. Sachs MB, Abbas PJ: **Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli**. *J Acoust Soc Am* 1974, **56**:1835–1847.
22. Palmer AR, Evans EF: **Cochlear fibre rate-intensity functions: no evidence for basilar membrane nonlinearities**. *Hear Res* 1980, **2**:319–326.
23. Dai H: **On measuring psychometric functions: a comparison of the constant- stimulus and adaptive up-down methods**. *J Acoust Soc Am* 1995, **98**:3135–3139.
24. Green DM, Swets JA: *Signal Detection Theory and Psychophysics*. Los Altos, California, USA: Peninsula Publishing; 1988.
25. Nizami L, Reimer JF, Jesteadt W: **The mid-level hump at 2 kHz**. *J Acoust Soc Am* 2002, **112**:642–653.
26. Nizami L: **The intensity-difference limen for 6.5 kHz: an even more severe departure from Weber's law**. *Percept Psychophys* 2006, **68**:1107–1112.
27. Zwillocki J, Maire F, Feldman AS, Rubin H: **On the effect of practice and motivation on the threshold of audibility**. *J Acoust Soc Am* 1958, **30**:254–262.
28. Lukaszewski JS, Elliott DN: **Auditory threshold as a function of forced-choice technique, feedback, and motivation**. *J Acoust Soc Am* 1962, **34**:223–228.
29. Loeb M, Dickson C: **Factors influencing the practice effect for auditory thresholds**. *J Acoust Soc Am* 1961, **33**:917–921.
30. Whitmore JK, Ermey HL, Williams PI: **Some results bearing on the stability of psychometric data**. *J Acoust Soc Am* 1968, **44**:370.

31. Weber DL, Moore BCJ: **Forward masking by sinusoidal and noise maskers.** *J Acoust Soc Am* 1981, **69**:1402–1409.
32. Zwicker E: **Dependence of post-masking on masker duration and its relation to temporal effects in loudness.** *J Acoust Soc Am* 1984, **75**:219–223.
33. Garcia-Perez MA: **A cautionary note on the use of the adaptive up-down method.** *J Acoust Soc Am* 2011, **130**:2098–2107.
34. Nizami L: **Dynamic range relations for auditory primary afferents.** *Hear Res* 2005, **208**:26–46.
35. Kohlrausch A, Puschel D, Alpehi H: **Temporal resolution and modulation analysis in models of the auditory system.** In *Speech research 10: the auditory processing of speech*. Edited by Schouten MEH. 1992:85–98.
36. Vaux DL: **Know when your numbers are significant.** *Nature* 2012, **492**:180–181.
37. Hartmann WH: *Signals, Sound, and Sensation*. New York: Springer-Verlag; 1998.
38. McGee JD: *Phase-locking as a frequency and intensity coding mechanism in auditory nerve fibers*. MS thesis: Creighton University; 1983.
39. Liberman MC: **Physiology of cochlear efferent and afferent neurons: direct comparisons in the same animal.** *Hear Res* 1988, **34**:179–192.
40. Palmer AR, Evans EF: **On the peripheral coding of the level of individual frequency components of complex sounds at high sound levels.** In *Hearing mechanisms and speech*. Edited by Creutzfeld O, Scheich H, Schreiner C. Heidelberg: Springer-Verlag; 1979:19–26.

doi:10.1140/epjnbp3

Cite this article as: Nizami: The Human Cochlear Mechanical Nonlinearity Inferred via Psychometric Functions. *EPJ Nonlinear Biomedical Physics* 2013 **1**:3.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com