NIH Public Access
Author Manuscript

Overshoot Measured Physiologically and Psychophysically in the Same Human Ears

Kyle P. Walsh, **Edward G. Pasanen**, and **Dennis McFadden**[a]
Department of Psychology and Center for Perceptual Systems, 1 University Station A8000, University of Texas, Austin, Texas 78712-0187, USA

Abstract

A nonlinear version of the stimulus-frequency otoacoustic emission (SFOAE) was measured using stimulus waveforms similar to those used for behavioral overshoot. Behaviorally, the seven listeners were as much as 11 dB worse at detecting a brief tonal signal (4.0 kHz, 10 ms in duration) when it occurred soon after the onset of a wideband masking noise (0.1 – 6.0 kHz; 400 ms in duration) than when it was delayed by about 200 ms, and the nonlinear SFOAE measure exhibited a similar effect. When either lowpass (0.1 – 3.8 kHz) or bandpass noise (3.8 – 4.2 kHz) was used instead of the wideband noise, the physiological and behavioral measures again were similar. When a highpass noise (4.2 – 6.0 kHz) was used, the physiological and behavioral measures both showed no overshoot-like effect for five of the subjects. The physiological response to the tone decayed slowly after the termination of the noise, much like the time course of resetting for behavioral overshoot. One subject exhibited no overshoot behaviorally even though his cochlear responses were like those of the other subjects. Overall, the evidence suggests that some basic characteristics of overshoot are obligatory consequences of cochlear function, as modulated by the olivocochlear efferent system.

Keywords

otoacoustic emissions; stimulus-frequency otoacoustic emissions; auditory masking; overshoot; temporal effect; auditory efferent system

I. INTRODUCTION

In all sensory systems, the incoming stimulus information is subjected to numerous stages of processing as it moves from the periphery to the cortex, where perception, consciousness, and behavioral responses presumably arise. An implicit, long-term goal of sensory neuroscience is to determine what modifications to the sensory stream are made at each successive stage of processing. In humans, the species for which the most is known about the behavioral aspects of sensory experience, there are obvious difficulties associated with gaining access to the successive physiological stages of processing. As a consequence,

[a]Corresponding Author: Dennis McFadden, Department of Psychology, Seay Building, University of Texas, 1 University Station A8000, Austin, TX 78712-0187, (512) 471-4324, (512) 471-5935 (fax), mcfadden@psy.utexas.edu.
Other Eddresses: kylewalsh@mail.utexas.edu pasanen@psy.utexas.edu

auditory science traditionally has relied on physiological measurements made on other species when developing explanations for human auditory experience. Here we report the use of a form of otoacoustic emission (OAE) capable of measuring the early stages of processing in humans. A strength of this measure is that it can be used with complex acoustic stimuli of the sort commonly employed to study various psychoacoustical tasks. The results suggest that these OAE measures may have potential to provide details about human cochlear processing in general and also about the individual differences in human cochlear processing that might be related to individual differences in human psychoacoustical performance.

OAEs are weak sounds produced inside the cochlea that propagate back out through the middle-ear system into the external ear canal where they can be recorded using small microphone systems (Kemp, 1978, 1979). There are several types of OAE (see Probst et al., 1991, for an early review). Of interest here is a version of the stimulus-frequency OAE (SFOAE), an OAE that is produced during the presentation of an acoustic stimulus. SFOAEs were studied early by Kemp and Chum (1980), Kemp (1980), Zwicker and Schloth (1984), Dallmayr (1987), and Lonsbury-Martin et al. (1990), and more recently by Guinan (e.g., Guinan, 2006; Guinan et al., 2003; Backus and Guinan, 2006), Keefe (e.g., Keefe, 1998; Schairer et al., 2003; Schairer and Keefe, 2005; Keefe et al., 2009), and others. Unlike some forms of OAE, SFOAEs have been measured successfully in non-human species as well as in humans (e.g., Goodman et al., 2003). To distinguish our measure from other versions of the SFOAE, we call our measure the nSFOAE (see Walsh et al., 2010).

Although many details have yet to be worked out, it is clear that the outer hair cells (OHCs) of the cochlea play an important role in the production of OAEs. When the basilar membrane moves up and down in response to an incoming sound, the OHCs are alternately depolarized and hyperpolarized, which in turn produces transformational changes in their length (contraction and elongation). As a consequence of this electromotility (Brownell et al., 1985), the displacement of the basilar membrane is greater than it is when the OHCs are non-functional, and hearing sensitivity is correspondingly better. Consequently, the OHCs are commonly viewed as forming an array of cochlear amplifiers (Davis, 1983). When the OHCs are absent or damaged along a segment of the basilar membrane, hearing sensitivity is reduced by about 40 dB (e.g., Smith et al., 1987).

The psychoacoustical phenomenon of primary interest here has been called overshoot or the temporal effect by different investigators (Zwicker, 1965a; McFadden, 1989; Bacon, 1990; Hicks and Bacon, 1992; Wright, 1995; Bacon et al., 2002; Strickland, 2001, 2004, 2008). When human listeners are asked to detect a brief tonal signal in the presence of a longer burst of masking noise, performance can differ substantially depending upon the relative timing of the masker and signal. Specifically, if the brief signal is presented shortly after the onset of the burst masking noise (short-delay condition), detectability can be as much as 10 – 20 dB worse than when the signal is presented 150 ms or so after the onset of the burst masker (long-delay condition). The magnitude of overshoot is largest when the tonal signal is high in frequency and only a few milliseconds in duration, and when the masking noise is broadband and relatively weak (Bacon, 1990; Overson et al., 1996). Early theorizing had placed the underlying mechanisms of overshoot in synaptic processes such as short-term depletion of neurotransmitter (e.g., Smith and Zwislocki, 1975; Smith, 1979; Westerman and Smith, 1984). More recent theorizing has suggested that the medial olivocochlear (MOC) efferent system acts to reduce the gain in the cochlear-amplification system and thus to make different cochlear input/output functions relevant for the short- and long-delay conditions (von Klitzing and Kohlrausch, 1994; Strickland, 2004, 2008; Keefe et al., 2003, 2009).

Supporting the idea that cochlear mechanics might contribute to the overshoot effect are some facts about how hearing loss affects performance in this task. Champlin and McFadden (1989) and McFadden and Champlin (1990) induced measurable hearing loss in normal-hearing ears by either exposures to intense sounds or administering high doses of salicylate (aspirin), respectively. Seemingly paradoxically, both of these manipulations *improved* detectability of the short-delay signal, but not the long-delay signal, and thereby diminished or eliminated the difference that defines overshoot. That is, temporary hearing loss made people *better* in the short-delay condition. When people with permanent hearing loss have been tested in overshoot tasks, they also typically have exhibited less of a difference between the short- and long-delay conditions than do normal-hearing people (e.g., Bacon and Takahashi, 1992; Turner and Doherty, 1997; Strickland and Krishnan, 2005), and, again, the difference has been smaller because detection in the short-delay condition was better than in normal-hearing people. That is, cochleas with either temporary or permanent damage to the cochlear-amplifier systems (Davis, 1983) curiously do better than normal cochleas at signaling tones presented soon after the onset of a burst masking noise.

The fact that both overshoot and OAEs appear to be dependent on a normal-functioning cochlea makes overshoot a natural choice for a study attempting to determine whether the cochlea makes a significant contribution to behavioral performance. If our nSFOAE response does exhibit parallels to overshoot measured behaviorally, it will suggest that cochlear function does play a significant role in overshoot, and it will make the nSFOAE a measure worthy of study with other psychoacoustical phenomena. A recent report suggested that cochlear function does *not* play a role in overshoot (Keefe et al., 2009), but our results, obtained with a different procedure, suggest the opposite. Some general properties of the nSFOAE response measured as part of this study are described in Walsh et al. (2010).

II. METHODS: GENERAL

A. Subjects

The data in this report were obtained from all four of the subjects used by Walsh et al. (2010) plus three additional subjects tested after that first group. The first group of subjects consisted of one female (aged 21) and three males (aged 26, 19, and 19). The second group consisted of three females (aged 27, 25, and 21). All subjects were screened for normal middle-ear function and normal hearing sensitivity (≤ 15 dB Hearing Level) in the right ear for the standard audiometric frequencies between 250 and 8000 Hz as measured by a clinical audiometric screening device (Auto Tymp 38, GSI/VIASYS, Inc., Madison, WI).

Two of the male subjects (JZ, NH) and three of the female subjects (EK, AB, YB) had 6-8 weeks of listening experience with various psychophysical tasks while serving in a related study, and one subject (author KW) had accumulated many hours of listening experience in a variety of psychophysical and physiological contexts. Female subject SC was the most inexperienced psychophysically; she was given several hours of training prior to formal data collection. All seven subjects had their OAEs measured at least once prior to this study. No subject had a spontaneous OAE (SOAE) stronger than −9.0 dB SPL any closer than 640 Hz to the 4.0-kHz tone used for most measurements here. Except for author KW, the subjects were paid for their participation. The female subjects were tested without regard to their menstrual cycle. Informed consent was obtained from all subjects prior to any testing, and the research protocol was approved by The University's Institutional Review Board.

III. METHODS: PHYSIOLOGICAL MEASURES

Our procedures for presenting the stimuli and extracting the nSFOAE response have been described in detail in a companion paper (Walsh et al., 2010), so they are only summarized

briefly here. We acknowledge in advance that the physiological procedures used were not perfectly parallel to the procedures used to measure overshoot behaviorally. Rather, the objective was to obtain a physiological measure of the functioning of the cochlea when presented with waveforms like those used behaviorally so that any parallels between physiology and behavior would be evident and so that the individual differences in the physiological measures could be compared with those anticipated to exist in the behavioral data (e.g., Zwicker, 1965a, b; Bacon and Liu, 2000).

For the OAE measures, individual subjects were seated in a comfortable reclining chair in a double-walled, sound-attenuated room. The subject relaxed alone in this room for 15 minutes prior to any data collection; an initialization period of this sort has been shown to enhance some forms of OAEs (Whitehead, 1991; McFadden and Pasanen, 1994). Prior to the initialization period, an Etymotic ER-10A microphone system (Etymotic, Elk Grove Village, IL) was placed snugly in the right external ear canal. For sound presentation, two Etymotic ER-2 earphones were attached to small plastic tubes that connected to the sound-delivery tubes passing through the microphone capsule into the external canal. The output of the ER-10A microphone was amplified 20 dB by the Etymotic preamplifier, and then passed to a custom-built amplifier/filter unit that highpass filtered the sound at 400 Hz and lowpass filtered it at 15 kHz. Digitizing of both the acoustic stimuli and the OAE responses was accomplished using a National Instruments board (PCI-MIO-16XE-10) installed in a Macintosh G4 computer; the sampling rate for both input and output was 50 kHz with 16-bit resolution. The acoustic stimuli were calibrated in a coupler (see Walsh et al., 2010, for details), and the waveforms produced by the computer were corrected according to that calibration prior to being presented to the earphones. Stimulus levels were measured in the ear canal and adjusted as necessary prior to each block of trials.

The sounds presented to the ear were synthesized digitally using the Macintosh G4 computer running custom-written LabVIEW® (National Instruments, Austin, Texas) software. For the physiological measurements, each trial consisted of three successive stimulus presentations (a triplet), and a block of trials consisted of at least 50 such trials. For the first stimulus presentation of each triplet, only one of the two Etymotic ER-2 earphones was activated; for the second presentation, only the other earphone was activated; and for the third presentation, both earphones were activated simultaneously. The electrical stimulus delivered to the individual earphones was always exactly the same in fine structure and level whether one or both earphones was being activated. The stimulus typically was tone-plus-noise for all three presentations of a triplet, but sometimes the tone was presented alone on all three presentations. Here, tone duration typically was 10 ms; in Walsh et al. (2010), it typically was 500 ms. The same noise sample was used for all presentations in all experimental sessions for all subjects.

The sound in the ear canal was recorded for all three presentations for each triplet. Those sounds consisted of the acoustic stimulus (and its reflected component owing to middle-ear impedance) plus whatever sound was being produced inside the cochlea in response to the acoustic stimulus. Because the waveform delivered to the earphones was always identical in fine structure, level, and starting phase whether one or both earphones was activated, if the cochlear response were strictly linear (and the middle-ear impedance constant), the instantaneous levels of the response to the third presentation would have been the exact sum of the responses to the first two presentations. For each triplet, the sounds obtained during each of the first two presentations were summed, and that sum was subtracted from the sound obtained during the third presentation of that triplet (a version of Keefe's "double-evoked" procedure; see Keefe, 1998). The result of this subtraction was a difference waveform containing only the nonlinear components of the SFOAE plus any residual nonlinearities in the measurement system.

The difference waveform extracted for each successive triplet was added to the accumulating nSFOAE average only if it met certain criteria, which were described in Walsh et al. (2010). Difference waveforms were obtained from 50 triplets in each block of trials and the sum of those individual difference waveforms was used to extract an estimate of the nSFOAE response to the sounds used for that block. Namely, the averaged difference waveform was filtered digitally with an 8th-order elliptical bandpass filter, 400 Hz in width and centered at 4.0 kHz. This filter was applied to successive 10-ms samples of the averaged waveform in 1-ms steps, the rms amplitude of the filter output was computed for each sample, and, after conversion to decibels sound-pressure level, the resulting succession of levels was taken as the estimate of the nSFOAE response. The analysis window was 20 ms for most of the data shown in Walsh et al. (2010); here it was reduced to 10 ms to better match the duration of the 10-ms tone used, with the necessary consequence that the data here show more moment-to-moment variability.

For some blocks of trials, the sound presented for all three presentations of all triplets was only a tone: the same 10-ms sample of a 4.0-kHz tone, gated with 5-ms rise/decay time (no steady-state portion) and fixed at some value of sound-pressure level. For other blocks, a sample of noise was presented along with the tone for all three presentations of each triplet. Typically, the noise had a bandwidth of 0.1 - 6.0 kHz, an overall level of about 63 dB SPL (corresponding to a level of 25 dB SPL in each 1-Hz band of the noise, hereafter called the spectrum level), a duration of 400 ms, and a rise/decay time of 2 ms for each presentation within each triplet. In some conditions, other bandwidths of the noise were used.

To increase the efficiency of data collection, two tone bursts were presented during each noise sample; for example, one delayed 50 ms from masker onset and one delayed 300 ms. The two tones had the same level and starting phase, and the time between the pair of tone bursts within a presentation was at least 195 ms. The segment of noise that occurred during tone presentation always was the same, irrespective of tone delay. This technique ensured that the interaction between each tone burst and the noise was identical. The blend point for the repeated segment of noise for the second tone presentation occurred midway between the two tone presentations. A 25-ms rise and decay was used for the blending. Always presenting the tone in the same fine structure of a frozen-noise sample was done in a behavioral overshoot context by von Klitzing and Kohlrausch (1994), and versions of this procedure have been used with SFOAEs by Keefe et al. (2003, 2009) and Walsh et al. (2008).

The various stimulus values of both tone and noise were choices known to give rise to substantial magnitudes of overshoot when used behaviorally (e.g., Zwicker, 1965a; Bacon, 1990; Overson et al., 1996; Strickland, 2001, 2004, 2008), and they were the same as those used for our behavioral measurements described below. Measurements also were made with the tone bursts presented after the offset of the noise by various time delays. In some conditions, each presentation of each triplet began with a 100-ms segment of the 4.0-kHz tone-alone to provide a reference against which subsequent changes could be compared (see Fig. 2).

Within the triplets, the silent period between the end of one noise presentation and the beginning of the next was about 500 ms in order to allow the auditory system to "reset" as fully as possible (see McFadden, 1989). (Actually, the silent times between successive presentations within triplets were chosen at random from the range 490 – 510 ms in an effort to reduce periodicity in the stimulus train.) *Between* the triplets, the silent period was as follows: 500 ms between the final presentation of triplet 1 and the first presentation of triplet 2, then 2000 ms between the final presentation of triplet 2 and the first presentation of triplet 3, and so on. The additional time after every other pair of triplets was required for real-time

calculations of the nSFOAE responses, comparison of the responses with the rejection criteria, updating the buffers, etc. Thus, the series of individual presentations was not perfectly regular at 500-ms intervals.[1]

IV. RESULTS: PHYSIOLOGICAL MEASURES

A. Tone-alone and noise-alone conditions

Figure 1 shows the nonlinear cochlear response when two 10-ms tone bursts of 60 dB each were presented during all three presentations of every triplet without the noise. For this demonstration, the presentations of the tones were delayed by 25 and 150 ms relative to the beginning of the recording period (for all other measurements, the minimum separation between tone pairs was 195 ms). Clearly, the nSFOAE response hovered around the noise floor of our measurement system (about −15 to −20 dB SPL) except during those moments when the tones were presented. The horizontal line in the figure reveals that the cochlear response to the long-delay tone was very similar to the response to the short-delay tone when only the tones were presented. The absolute magnitudes of those responses did vary directly with the sound-pressure level of the tone bursts (see Walsh et al., 2010, Fig. 1).

It is important to be explicit about what is plotted in Fig. 1. As noted above, the averaged response waveform was analyzed in 10-ms windows advanced in 1-ms steps. For each step, the rms amplitude was calculated for a 400-Hz bandwidth centered on 4.0 kHz. That succession of amplitudes, expressed as decibels sound-pressure level, is the succession of data points plotted in Fig. 1 and in the other nSFOAE figures here. Note that the abscissa value of each data point marks the *beginning* of a 10-ms window. Thus, the data point at 20 ms represents the strength of the nSFOAE response from 20 to 30 ms. Careful examination of Fig. 1 reveals that the responses begin to rise prior to the 25- and 150-ms delays at which the tonal stimulus actually was presented. This is expected because the beginning of the response should contribute first to the very end of a time window whose beginning precedes the stimulus presentation by less than 10 ms. What is interesting is that the peak of the response is delayed by about 3 ms from the time expected if our response were simply the tonal stimulus itself. For example, the peak responses in Fig. 1 occur at about 28 and 153 ms, not at 25 and 150 ms as they should if the 10-ms time windows were seeing the acoustic stimulus instead of a physiological response. This approximately 3-ms delay corresponds favorably with estimates of the round-trip travel time from the external ear canal to the 4.0-kHz location along the basilar membrane (e.g., Shera and Guinan, 2003; cf. Schairer et al., 2006), but the latency measures that emerge from our procedure are difficult to interpret (see Results section IV E below).

When a single tone of 400 ms duration was presented alone (no noise) instead of short tone bursts, the magnitude of the nSFOAE response rose immediately to essentially the same value as seen for a tone burst of that same sound-pressure level, and it remained at that value for the duration of the presentation. For both short- and long-duration tones, the absolute magnitude of the nSFOAE response does vary directly with the level of the tone (long-tone data are shown in Walsh et al., 2010, Fig. 1).

When a wideband noise was presented alone, the average magnitude of the nSFOAE response increased gradually over the time course of the presentation, at least for some

---

[1]Prior to the collection of all the data described here, an initial full set of data was collected for the first crew of four subjects, using stimulus waveforms that were mis-calibrated slightly. Because of the mis-calibration, the noise bands for some conditions were as weak as 21 dB SPL spectrum level (instead of 25), and the tone was as strong as 66 dB (instead of 60). Even so, for those four subjects, all of the basic findings were the same as described here. The second crew of three subjects was tested after the calibration error was corrected.

subjects and for moderate noise levels (spectrum levels of 20 and 25 dB; see Fig. 3 below, and Walsh et al., 2010, Fig. 2). At the strongest and weakest noise levels tested (spectrum levels of 35 and 15 dB, respectively), the increase was small or absent for most subjects.

B. Tone-plus-noise conditions

When the brief tone bursts were presented during a longer noise burst, the results were markedly different from either the tone-alone or noise-alone responses. Specifically, as the onset delay of the tone burst was increased relative to the onset of the noise, the maximum magnitude of the nSFOAE response at the frequency of the tone increased dramatically during approximately the first 100 ms of delay and then stayed essentially constant for the remainder of the noise duration. At asymptote, the magnitude of the nSFOAE response was about 6 - 19 dB stronger than for a tone of that same level presented alone. This gradually increasing nSFOAE response was essentially the same whether tone bursts or a single long-duration tone were presented (see Fig. 2; also see Walsh et al., 2010, Fig. 10, for a similar comparison). Apparently, whatever mechanism is producing this gradual rise in nSFOAE response at the frequency of the tone is indifferent to the duration of that tone. Pragmatically, this means that short probe tones can be used just as well as long-duration tones to measure the state of this dynamic process.

Fig. 3 shows a comparison of the nSFOAE responses obtained with tone-alone and tone-plus-noise. For this example, every presentation of every triplet involved two tones with a separation of 125 ms between them whether one or both earphones was being activated. The solid symbols are the data collected with tones-alone also shown in Fig. 1; the open symbols are the data obtained when the two tones were presented simultaneously with a 400-ms, wideband noise on every presentation of every triplet. Clearly, when the tone was presented with the wideband noise, the nSFOAE response was substantially stronger for the 150-ms delay than for the 25-ms delay; also, the response at the 25-ms delay was highly similar with or without the noise. The gradual increase in nSFOAE response in the time between the short- and long-delay tones is in accord with the response seen to noise-alone with moderate noise levels (see Walsh et al., 2010, Fig. 2).

The fact that the nSFOAE response to the tone having 25-ms delay was highly similar whether the noise was present or not suggests that the strengths of our nSFOAE responses were determined primarily by the tone, not the (weaker) noise. In another set of measurements (using long-duration tones), the level of the tone was held constant while the level of the noise was manipulated across blocks of trials. The result was that the maximum magnitudes of the dynamic nSFOAE response changed relatively little over a 20-dB change in the spectrum level of the wideband noise (see Walsh et al., 2010, Fig. 8). This outcome also suggests that the existence of a rising, dynamic response depends crucially upon the presence of the noise, but the magnitudes of the nSFOAE responses are determined primarily by the tone.

In one regard, the data in Fig. 3 are not strictly representative of the data seen in other subjects. Note in Fig. 3 that the peaks in the nSFOAE responses exhibit a slight shift toward shorter time values in the tone-plus-noise condition (a latency shift) compared to the tone-alone condition. This pattern has been seen in other subjects and situations, but it is not universal across subjects and conditions. Sometimes only the response peak for the long-delay tone shifts toward smaller time values; sometimes the response peaks for neither show any time shift. The pattern seen is not solely a function of the individual ear; the use of a different sample of synthesized noise in the same ear can change the pattern, suggesting that the tone and noise can interact in a way that affects the timing of the peak in the nSFOAE response. Noise level also can matter. In some ears with some noise samples, increases in level of the wideband noise can lead to increasing latency shifts; in others, the presence of

the noise causes a latency shift compared to tone-alone, but that latency shift does not change with increasing noise level. Additional research will be required before the issue of latency shift is resolved. We return below to this latency shift when discussing the possible contribution of the middle-ear reflex to our nSFOAE responses.

Data like those in Fig. 3 were collected for all seven subjects using numerous pairs of time delays for the tone; the results are shown in Fig. 4. These functions all rise gradually over the course of tens of milliseconds and then either asymptote or decline slightly. However, the subjects exhibited marked individual differences (also see Lilaonitkul and Guinan, 2009a), with subjects JZ and EK showing maximum responses that were about 16 and 19 dB above their initial values, respectively, and subject SC showing a maximum response only about 6 dB above her initial value. When long-duration tones are used, there often is a hesitation of about 20 - 25 ms after noise onset before the rising, dynamic response begins (Backus and Guinan, 2006;Walsh et al., 2010); similar hesitations can be seen for short tones in the data for some of the subjects in Fig. 4. All of the subjects tested using long-duration tones did show a hesitation of approximately 25 ms. An exponential function of the form $y = a(1 - e^{bx}) + c$ was fitted to the data for each subject beginning 20 ms after noise onset; those fits and the corresponding time constants are shown in each panel of Fig. 4. As can be seen, most of the time constants are relatively short (compare Backus and Guinan, 2006), and they would be even shorter if an analysis window longer than 10 ms had been used (see Walsh et al., 2010). Time constants for gradual changes in OAEs following noise onset also have been studied by Maison et al. (2001),Kim et al. (2001), and Bassim et al. (2003); for a review see Backus and Guinan (2006).

The data in Figs. 3 and 4 are interesting in the context of auditory masking because the magnitude of this nonlinear cochlear response to a 10-ms tone burst depended upon the timing of the tone relative to noise onset. Because weaker response magnitudes were observed for brief tones presented near the onset of the simultaneous wideband noise than for tones presented with longer onset delays, the directionality of effect is the same as for overshoot measured psychophysically. That is, an effect seen behaviorally in auditory masking experiments also is seen in the nSFOAE response from the cochlea. This rough parallel between the nSFOAE response and psychophysical measures of auditory masking led us to compare the two domains of measurement more closely. If this physiological measure were to behave similarly to behavioral measures obtained from the same individual ears, then the nSFOAE response would be established as a potentially valuable window on the early stages of processing in certain auditory masking tasks.

C. Spectral characteristics of the noise

As reported previously (Walsh et al., 2010), the rising nSFOAE response can be activated by just the low-frequency components of the wideband noise. However, it cannot be activated by just the high-frequency components or just the components centered on the frequency of the tone; rather, with highpass and bandpass noises, the nSFOAE was weak and approximately constant at about the level of the response to tone-alone. That previous demonstration (Walsh et al., 2010) used long-duration tones. Of interest here are the magnitudes of the nSFOAE responses to our brief tone bursts presented with both short and long delays when the noise has various bandwidth configurations. Accordingly, nSFOAE data were collected with 10-ms, 4.0-kHz tones having delays of 5 and 200 ms from the onset of the 400-ms sample of noise. The bandwidth of the noise was either 0.1 – 3.8 kHz (lowpass), 3.8 – 4.2 kHz (bandpass), or 4.2 – 6.0 kHz (highpass); for all bandwidths, the spectrum level was held constant at 25 dB.

Similar to the outcome using long-duration tones, the overshoot-like difference in nSFOAE magnitude for short- and long-delay tones was largest when the noise was wideband, but it

also was relatively large when the noise was lowpass filtered. By comparison, the magnitude of the nSFOAE response was about the same for the short- and long-delay tones when the noise was either highpass filtered above the tone or bandpass filtered around the tone. In Fig. 5 are shown the differences in nSFOAE magnitude for the short- and long-delay tones. The heights of the bars designate the mean differences between the short- and long-delay conditions across the seven subjects, and the data points to the immediate left of each bar show those differences for each individual subject. Supplementary measurements in which the overall levels of the noise bands were held constant at 63 dB SPL also showed no rising, dynamic response for the bandpass and highpass noise bands even though this corresponded to increases in spectrum level of about 12 and 5 dB, respectively. Nelson and Young (2009) observed a similar asymmetry of spectral effect on neural responses from the inferior colliculus of marmoset monkeys using stimulus waveforms related to those used here.

D. Recovery following noise termination

In Fig. 4, the nSFOAE responses were still at or near their maximum values when the 400-ms noise burst terminated. A question of interest is how rapidly the response returns to baseline following the termination of the noise. To answer that question, 10-ms tone bursts were presented at various times after the *offset* of a 200-ms wideband noise. This duration was chosen for the noise because the nSFOAE responses for all subjects had reached asymptote by this time, and data collection was more efficient using 200-ms noise samples than 400-ms samples. Again, for efficiency, two tones with a separation of at least 200 ms were presented during each triplet in each block of trials.

The results for the three subjects tested revealed that recovery was quite slow (see Fig. 6). To return to the tone-alone baseline required more than 600 ms of silence for all three subjects. Thus, the time required for the nSFOAE response to return to baseline magnitude was considerably longer than the time required to rise from baseline to maximum magnitude (see Fig. 4). (Related measurements collected with a long-duration tone and varying noise durations are shown in Walsh et al., 2010, Fig. 9.) Goodman and Keefe (2006) also have reported long recovery times following the offset of a wideband noise. Note that the data in Fig. 6 suggest that our use of approximately 500 ms between the successive presentations in each triplet (see section III above) was not quite adequate to allow the system to return completely to its initial state in some subjects. To the extent that 500 ms was inadequate for complete resetting in any subject, the differences in the nSFOAE responses for short- and long-delay tones (the overshoot-like effect) are likely to be underestimates. When the tone was 66 dB, the recovery times were generally shorter and the variability across subjects greater.

E. Middle-ear reflex

As noted, the parameters of the noise and tones used here were chosen on the basis of their ability to produce overshoot behaviorally. Unfortunately, while it may be close to optimal for obtaining overshoot behaviorally (Bacon, 1990; Overson et al., 1996), a wideband noise having an overall level of 63 dB (first two presentations in each triplet) or 69 dB (third presentation in each triplet) also may have the ability to activate the middle-ear reflex (MER) in some subjects, and measures like the nSFOAE may be affected accordingly (Guinan, 2006; Guinan et al., 2003; Backus and Guinan, 2006). Various facts suggest that the MER was not a significant factor in the present measurements; see Walsh et al. (2010) for additional discussion.

1. The MER primarily attenuates frequencies below about 1.0 kHz (Dallos, 1973; Goodman and Keefe, 2006; Schairer et al., 2007), and the tone used here was 4.0 kHz.

2. The rising, dynamic segment of the nSFOAE response was strongly dependent upon the bandwidth of the noise presented. Bands of noise centered on the 4.0-kHz tone or high-passed above the tone were completely ineffective at producing a rising, dynamic response to the tone, even when the level of the noise was as much as 12 dB greater than used to collect the data in Fig. 5. However, a band of noise low-passed below the tone did produce a dynamic segment (Fig. 5). In contrast, the MER can be activated by sounds all across the spectrum, with high frequencies generally being more effective than low frequencies (Dallos, 1973).

3. The onset latency for the MER is in the range of 70 - 100 milliseconds depending upon the level of the stimulus (Dallos, 1973, Fig. 7.13; Church and Cudahy, 1984; Goodman and Keefe, 2006), but the rising, dynamic phase of our nSFOAE response begins about 20 - 25 ms after the onset of a relatively weak wideband noise.

4. For both the tone-alone and tone-plus-noise presentations shown in Figs. 1 and 3, the peaks of the nSFOAE responses were delayed about 3 ms compared to the peaks of the acoustic stimuli. As noted, these delay values varied depending upon the subject, the particular sample of noise, and the level of the noise, but never were they zero, and never did they suggest contamination by an MER-induced reflection having zero latency.[2]

5. For subjects NH and KW, the frequency of the tone was varied in 5-Hz steps around 4.0 kHz in the presence of a wideband noise, and the phase of the nSFOAE response shifted systematically as it should if the response originates within the cochlea (see Guinan et al., 2003).

V. METHODS: BEHAVIORAL MEASURES

For the behavioral measurements, subjects were seated in individual listening booths in a large, double-walled, sound-attenuated room. Each subject was provided with an array of indicator lights that marked the various intervals of each trial: warning interval and light (350 ms), pause (500 ms), first observation interval and light (400 ms), pause (500 ms), second observation interval and light (400 ms), response interval (1000 ms), and feedback interval and light (350 ms). The subjects were provided with TDH-39 headphones (300 ohms) mounted in circumaural cushions. Sounds were presented only to the right earphone. The first group of four subjects was tested simultaneously, as was the second group of three subjects, with the exception of a few individual make-up sessions.

Behavioral data were collected using an adaptive (three-up, one down), two-interval forced-choice (2IFC) procedure (Levitt, 1971). Both observation intervals of each trial contained a 400-ms burst of masking noise (typically 0.1 – 6.0 kHz). The 10-ms tonal signal was presented in one of the two observation intervals at random, and the subject pressed one of two response keys to indicate which interval contained the signal. The onset of the signal followed the onset of the 400-ms masker by one of 10 delay values, ranging from 5 ms to

---

[2]Although its presence and its magnitude suggest that this 3-ms delay is simply a consequence of the round-trip travel time from the external ear canal to the 4.0-kHz location along the basilar membrane, one needs to be cautious when interpreting this number. The reason is that the nSFOAE is obtained by taking the difference between the summed one-earphone presentations and the two-earphone presentation. If the SFOAEs for the former and the latter had slightly different propagation delays, the peak in the resultant nSFOAE waveform would necessarily be at a value intermediate to those individual peaks. If the difference in the propagation delays were large, the envelope of the response would not match that of the gated stimulus; it could be too broad, its peak could be too flat, or it could have two shallow peaks. In an extreme case, if the MER were activated by the two-earphone presentation but not by the two single-earphone presentations, the envelope of the response could contain an additional, short-latency peak owing to reflections from the tympanic membrane. Although we never have observed envelope distortions of this sort, we caution the reader against interpreting our 3-ms delays as veridical estimates of round-trip travel time.

385 ms. The value of signal delay was varied across blocks of trials but was constant for all trials of each block. Of course, only one signal was presented per observation interval, unlike the physiological conditions. In some blocks of trials, the signal was presented after the termination of the noise (forward masking) by time delays ranging from 2 to 20 ms. The offset of the masker burst presented in the first observation interval of each trial was separated from the onset of the masker burst of the second observation interval by 500 ms, the same separation as used for the physiological measurements; this long interval was desirable in order to allow the auditory system to "reset" following the first noise burst (see McFadden, 1989).

In accord with Strickland (2004), the tonal signal was fixed in level for a block of trials (typically at 60 dB SPL), and the level of the *masker* was adjusted on a trial-by-trial basis. After three consecutive correct decisions, the strength of the masker was increased by 2 dB, and after each incorrect decision, it was decreased by 2 dB. At the end of each 50-trial block, the first two or three reversals were discarded and the remaining even number of reversals was averaged to produce an estimate of the masker level required by that subject for 79% correct detections of the signal. Blocks having fewer than 45 responses, fewer than three reversals, or a standard deviation of the reversals greater than 3.5 dB were discarded. At least three 50-trial blocks were collected for each condition. During some blocks of trials, detectability was measured for the 10-ms, 4.0-kHz tone in the quiet. The masker and signal waveforms were generated with 16-bit resolution at a sampling rate of 50 kHz using a digital/analog converter (PCI-4451, National Instruments, Austin, TX) installed in a Macintosh G4 computer. The same computer was responsible for presenting the stimuli and collecting the data. The level of the masking noise was adjusted adaptively for each subject individually using 12-bit programmable attenuators (Charybdis, Model D).

Just as for the physiological measurements, the masking noise typically had a bandwidth of 0.1 – 6.0 kHz. The noise and tone were gated with cosine-squared gating functions having rise/decay times of 2 ms and 5 ms, respectively. Thus, the signal had no steady-state segment.

VI. RESULTS: BEHAVIORAL MEASURES

A. Signal delay

Only six of the seven subjects exhibited overshoot psychophysically with the stimulus parameters used here. The basic data are shown in Fig. 7. As can be seen, all subjects except for author KW required weaker levels of the masker (poorer detection of the fixed-level signal) when the onset of the signal was close to the onset of the 400-ms masking noise than when signal onset was delayed by 100 ms or more. When performance was compared for the conditions with a 5-ms delay and a 200-ms delay, the overshoot magnitudes were 0.4, 6.4, 11.2, 9.7, 4.8, 5.5, and 7.1 for subjects KW, NH, SC, JZ, EK, AB, and YB, respectively. These values of overshoot are, for some reason, smaller than often reported at this signal frequency (e.g., Strickland, 2008), even though the masker levels for all subjects were generally in the range that produces the largest overshoot (Bacon, 1990;Overson et al., 1996).

Each data set in Fig. 7 was fitted with the exponential function used for the physiological data in Fig. 4, and the resulting time constants again are shown in each panel. For these fits, no 20-ms hesitation was implemented. There are both intriguing similarities and marked discrepancies between the individual physiological and psychophysical data plotted in Figs. 4 and 7, respectively. For example, subject JZ has rapidly rising functions for both physiology and psychophysics and subject AB has slowly rising functions for both, whereas subject KW has typical-looking data in Fig. 4 but decidedly atypical data in Fig. 7.

For comparison with these masking results, the signal level necessary for detection in the quiet was similar across the six subjects showing overshoot (mean = 27.8 dB SPL, SD = 2.3), with the greatest sensitivity (24.9 dB SPL; subject SC) being about 6 dB better than the worst (31.0 dB; subject NH). The sensitivity of subject KW was 26.2 dB, seemingly ruling out hearing loss as an explanation for his lack of overshoot.

B. Spectral characteristics of the masker

As shown in Fig. 5 (also see Walsh et al., 2010,Fig. 6), the nSFOAE response to tone-plus-noise is highly dependent upon the spectral characteristics of the noise. Specifically, it is primarily the low-frequency components of the noise that activate the rising nSFOAE response to the tone. Past research has differed about how bandwidth manipulations affect overshoot measured behaviorally (e.g., McFadden, 1989;Carlyon, 1989;Hicks and Bacon, 1992;Bacon et al., 2002;Strickland, 2004), so in order to compare our subjects' physiological and psychophysical measures, the same bandwidth manipulations made for the physiological conditions were made for the psychophysical conditions. To be specific, psychophysical data were collected for 4.0-kHz signals having short (5 ms) or long (200 ms) delays after the onset of a 400-ms masking noise that was either lowpass filtered below the signal (0.1 – 3.8 kHz), bandpass filtered around the frequency of the signal (3.8 – 4.2 kHz), or highpass filtered above the signal (4.2 – 6.0 kHz). As before, the level of the 10-ms signal was fixed (60 dB) and the level of the masker was varied adaptively.

The psychophysical data obtained using the various noise bandwidths are shown in Fig. 8. The heights of the bars designate the mean differences between the short- and long-delay conditions (the overshoot), and the data points adjacent to each bar show the overshoot values for each individual subject. On average, the wideband and lowpass noises led to similar magnitudes of overshoot (about 6.5 dB when subject KW was included in the calculations and about 7.2 - 7.5 dB when he was not), while the bandpass noise produced no overshoot on average. The highpass noise did produce an average overshoot of about 3.0 dB, but that mean was inflated by two subjects (with them omitted, the average overshoot for highpass noise was 0.7 dB). The absence of overshoot with the bandpass noise is in accord with past research (Zwicker, 1965b;McFadden, 1989); however, some past research (Schmidt and Zwicker, 1991;Hicks and Bacon, 1992) had suggested that the high-frequency components of the noise are more critical to overshoot than the current data suggest.

For our subjects, then, wideband and lowpass noises were more successful at producing differences between the short- and long-delay conditions than were the bandpass and highpass noises, *and* that was true both physiologically and psychophysically. For the highpass noise, the average difference between short- and long-delay conditions was essentially zero physiologically and was non-zero psychophysically; however, that discrepancy in the means was not representative of the individual data for the majority of the subjects. The psychophysical data (Fig. 8) did exhibit more variability across subjects than the physiological data (Fig. 5), but the overall patterns of results were generally similar in the two domains.

C. Forward masking

Forward masking is reported for the 10-ms, 4.0-kHz signal for three values of delay (2, 10, and 20 ms) following the offset of the masking noise (0.1 – 6.0 kHz in width and 200 ms in duration). The signal level was fixed at 45 dB SPL, and the level of the noise was varied adaptively to estimate the level required for 79% correct detections. This weak signal level was used because preliminary measurements indicated that the 60-dB signal used for the other tasks required masker levels higher than the maximum value judged to be free from possible effects of temporary hearing loss (90 dB overall). Even with this weaker signal, all

subjects required this maximum masker level for an additional signal delay of 40 ms, so data collection was cancelled for that condition. The results for the other three delays are shown in Fig. 9.

In accord with a substantial literature on forward masking (e.g., Jesteadt et al., 1982), the signal was monotonically more detectable (forward masking decreased) as signal delay was increased. Subject SC was clearly best at detection in this task (she required the strongest masker levels); she also had the best absolute sensitivity for the 4.0-kHz, the most overshoot (Fig. 7), the most SOAEs in the test ear, and the strongest nSFOAE response to the 4.0-kHz tone alone. When the data were fitted with straight lines, the slopes were 1.9, 1.3, 1.4, and 1.6 dB/ms for subjects SC, NH, KW, and JZ, respectively. As discussed below, the time course of recovery for these data is markedly faster than the decay observed for the nSFOAE response (Fig. 6).

VII. DISCUSSION

A. Comparison of physiological and psychophysical outcomes

In order to evaluate how useful our nSFOAE measurement might prove to be, it is helpful to compare the outcomes obtained physiologically and psychophysically when the same stimulus manipulations were made.

In support of the value of the nSFOAE measure, all seven subjects showed a stronger nSFOAE response in the tone-plus-wideband-noise condition when the onset of the brief tone was delayed by 100 ms or more after the onset of the noise than when its onset followed the onset of the noise by just a few milliseconds (Fig. 4). This dynamic pattern of response is reminiscent of the psychophysical phenomenon of overshoot; it is exactly the pattern of response that would be expected if a stronger nSFOAE response is related to a stronger cochlear response to the tone than to the noise. Furthermore, the time constants estimated for the dynamic nSFOAE response (about 26 to 128 ms) were reasonably similar to those estimated from the psychophysical overshoot data (about 23 to 142 ms) obtained from the same subjects.[3]

These facts might lead one to conclude that overshoot is an obligatory consequence of the mechanics of the human cochlea; short-delay tones simply elicit weaker cochlear responses than long-delay tones, and that weakness is preserved through subsequent stages of neural processing, ultimately resulting in the overshoot measured behaviorally. Contradicting this attractive view is the fact that one subject (author KW) exhibited a perfectly typical rising physiological response with the wideband noise even though he exhibited no overshoot psychophysically with that same noise. If the nSFOAE response were obligatorily related to the mechanisms underlying overshoot, then it is not clear how a subject could manage to overcome the necessity to hear a short-delay signal less well than a long-delay signal. That would seem to require a neutralization at some higher neural level of the disadvantage short-delay signals have when leaving the cochlea, and if such neutralization were possible in the wideband condition, then why not in the lowpass condition as well, and why not in all subjects?

The nSFOAE response to tone-plus-noise seems to be based largely on the frequency components lying below the tone. The wideband noise did lead to the largest difference in

---

[3]The similarity of these time constants from the two domains needs to be interpreted with care because neither the procedures used to collect the physiological and psychophysical data nor the fitting procedures were perfectly parallel. For example, the noise level was variable for the psychophysical measures and fixed for the physiological measures, and no hesitation was implemented for the psychophysical fits.

the nSFOAE responses to the short- and long-delay tones, but a lowpass noise also led to a difference of considerable magnitude (Fig. 5). In contrast, the bandpass noise was unable to activate the dynamic phase of the nSFOAE. The behavioral data were in agreement with all these physiological outcomes: psychophysically, overshoot also was largest for the wideband noise, reasonably large with the lowpass noise, and nonexistent with the bandpass noise (Fig. 8). Countering all these agreements between the psychophysical and physiological data is the result that there was no dynamic phase to the nSFOAE response for any subject when the noise was highpass, but that same noise did lead to considerable overshoot behaviorally for two of the seven subjects.

The nSFOAE response to tone-plus-noise typically shows a hesitation of about 20 - 25 ms before exhibiting the characteristic rising, dynamic response (also see Backus and Guinan, 2006). We always see this hesitation when using long-duration tones, but see it only sometimes when using short-duration tones. None of our behavioral data exhibited a hesitation, but fine-grained manipulations of signal delay were not made. To our knowledge, psychophysicists studying overshoot in the past also have not used multiple signal delays shorter than 25 ms, so it is still possible that behavioral hesitation may exist.

Thus, there were several similarities and some differences in the physiological and psychophysical responses to manipulations of noise bandwidth. Outcomes of this sort are plausible if one makes the reasonable assumption that the cochlea can impose obligatory characteristics on behavioral overshoot and that higher neural centers also can contribute characteristics (e.g., overshoot with highpass noise) before behavioral responses are determined. However, even under this interpretation, the dissociation between physiology and psychophysics for subject KW remains a problem because it seems to require that degraded performance at the cochlear level can somehow be restored at a higher stage of processing.

As noted above, the nSFOAE response clearly is an indirect and imperfect measure of the cochlear processes contributing to any psychophysical phenomenon such as overshoot. The procedures used to obtain the nSFOAE response are necessarily different from those used psychophysically, and the nSFOAE procedures limit us to seeing only any nonlinear component that might exist in the cochlear response. So, the best that ever can be said is that the nSFOAE response is correlated (or not) with the psychophysical results. Nonetheless, the nSFOAE does appear to carry information about cochlear effects that may be relevant to psychophysics. There is precedent for using different procedures physiologically and psychophysically to study an auditory phenomenon; for example, the common physiological procedure for studying lateral suppression involves simultaneous presentation of signal and suppressor (e.g., Sachs and Kiang, 1968; Keefe et al., 2008), whereas the most common psychophysical procedure involves forward masking (Shannon, 1976).

Forward Masking and Resetting—The nSFOAE response to brief tones presented after the offset of a 200-ms wideband noise gives a first impression of behaving like forward masking in that the response remains higher than the response to tone-alone for many tens of milliseconds. But a comparison with psychophysical forward-masking data collected on the same ears revealed that the time courses were quite different for physiology and psychophysics. It is unclear how best to compare the data from the two domains, but a simple procedure is to ask how much the time delay had to increase for the criterion measure to change by a fixed amount. For detectability of the signal to improve 10 dB in the psychophysical task required about 8 – 10 ms of additional tone delay (Fig. 9), whereas for the magnitude of the nSFOAE to decline 10 dB required hundreds of milliseconds depending upon the subject (Fig. 6). (This comparison probably underestimates the difference because the physiological data were collected with a 45-dB tone and the

psychophysical data were collected with a 60-dB tone.) Thus, the nSFOAE response appears not to be closely related to the mechanisms underlying auditory forward masking measured psychophysically.

However, there is a psychophysical outcome called "resetting" (McFadden, 1989) that does appear to be related to the post-offset persistence of the nSFOAE response, and furthermore is a basic characteristic of behavioral overshoot. If a brief temporal gap is inserted into long-duration noise just before the presentation of a brief tonal signal, performance continues to be as good as is to be expected with a signal presented long after the onset of a noise. However, as the duration of that temporal gap is increased, performance inevitably begins to approach what is seen when a signal is presented soon after the onset of a noise (McFadden, 1989; Overson et al., 1996). That is, as the gap duration increases, signal detectability changes from being like a long-delay condition to being like a short-delay condition. This resetting process proceeds quite slowly, however, with gap durations upwards of 300 ms being required to move detectability from long-delay to short-delay levels. These long resetting times are reminiscent of the long course of recovery seen for the nSFOAE response following noise offset (Fig. 6), and we suggest that this comparison is more appropriate than a comparison with behavioral forward masking. Strengthening this argument is the fact that resetting of overshoot is controlled by frequency components remote from the signal frequency (see McFadden, 1989; Overson et al., 1996), just as is true for the nSFOAE (Figs. 5 and 8). Thus, there appears to be yet another important point of agreement between the nSFOAE response and overshoot measured behaviorally.

B. Other attempts to study overshoot with OAEs

Prior to our developing the procedures used to extract the perstimulatory nSFOAE response described here, other OAE procedures were used in the search for possible concomitants to overshoot in cochlear mechanics. Specifically, versions of transient-evoked OAEs (TEOAEs) were developed that allowed extraction of the echoes produced by short tone bursts presented at two different delays after the onset of a noise (Walsh et al., 2008). In order to make the echoes produced by the tone more accessible for averaging, the noise bursts used had spectral notches at the frequency of the tone, and these synthesized noise bursts also had identical fine structure both across trials and during the presentations of the short- and the long-delay tones within a trial. The analysis process was such that both the linear and nonlinear components of the echo-like response to the tone burst were extracted. Psychophysical data also were collected using essentially the same waveforms as used for the OAE measurements. For both psychophysics and physiology, the spectrum level of the noise was 25 dB, which is within the range recommended for maximum overshoot magnitude (Bacon, 1990; Overson et al., 1996).

In accord with Strickland (2004), all of the eight normal-hearing subjects tested in that earlier study had moderate or considerable behavioral overshoot for both the notched-noise and wideband conditions. However, the TEOAE measures contained no evidence that the cochlear response was weaker for the short-delay tone than for the long-delay tone. The implication is that the TEOAE response (a post-stimulatory measure) is not closely associated with whatever mechanisms underlie human detection of a brief tone in those listening conditions that lead to overshoot. Eventually the difference between the TEOAE and nSFOAE results may prove to be attributable to the TEOAE containing both linear and nonlinear components and the nSFOAE containing only nonlinear components.

Keefe et al. (2009) recently have reported work having exactly the same motivation and rationale as this work, an attempt to use SFOAEs to measure possible cochlear concomitants to behavioral overshoot. They also used the same stimuli, the same ears, and similar procedures for their physiological and psychophysical measures. (None of this work was

known to us when planning or conducting our studies using TEOAEs or nSFOAEs.) Many details of the stimuli and procedures were similar in the two studies; however, the outcomes are markedly different. Keefe et al. (2009) reported no overshoot-like difference in their SFOAE measure for the short- and long-delay conditions. One, presumably quite important, difference between the two studies was that the triplet sequence used by Keefe et al. was: tone-plus-noise, suppressor-tone-alone, tone-plus-noise-plus-suppressor. In that procedure the level of the noise never was doubled by simultaneous presentation on the two earphones, as it is for our nSFOAE. Also, the physiological procedure used by Keefe et al. estimated the strength of the tone necessary for the resulting SFOAE to be just detectable from the noise floor of their recording system (a "threshold" measure) whereas our procedure measures the strength of the nSFOAE produced by a fixed level of the tone and/or noise (a "suprathreshold" measure). The tonal suppressor used by Keefe et al. for their physiological measures apparently was omitted from their psychophysical task, but we have conducted some pilot work that suggests that the presence of a suppressor tone does not noticeably affect the magnitude of behavioral overshoot.

In a previous report, Keefe et al. (2003) did observe an overshoot-like difference between the short- and long-delay conditions when using an earlier version of their SFOAE procedure. However, like here, some of the overshoot-like differences Keefe et al. (2003) observed with that earlier SFOAE measure did not behave exactly the same as do behavioral measures of overshoot (obtained on different subjects) as stimulus parameters were varied. Namely, the greatest difference between short- and long-delay tones was seen when the spectrum level of the noise was high, not moderate (compare Bacon, 1990; Overson et al., 1996).

So, the history of searching for overshoot-like effects using OAE responses from the cochlea is mixed. Depending upon seemingly subtle differences in procedure and stimuli, one either can find evidence for a weaker cochlear response for short-delay tones than for long-delay tones or not. What is needed is insight into why subtle procedural differences matter so much and which procedure(s) are best for which questions.

C. Efferent system and cochlear input/output functions

The procedures and outcomes reported here clearly are related to those described by Guinan (e.g., Guinan et al., 2003; Guinan, 2006; Backus and Guinan, 2006), so it is natural to presume that similar underlying mechanisms are involved in the two sets of measurements. Specifically, using a different form of SFOAE, Guinan and colleagues also have reported a dynamic response when a tone is presented along with a wideband noise, and they have offered considerable evidence and compelling arguments that at least the early segment of that dynamic response is attributable to the actions of the medial olivocochlear (MOC) efferent system acting on elements in the cochlea. If we apply Guinan's interpretation to the data presented here, it means that the behavioral phenomenon of overshoot is (largely) attributable to the actions of the MOC system, which is exactly the suggestion made by von Klitzing and Kohlrausch (1994) and further elaborated by Strickland (2004, 2008) for behavioral overshoot and applied to SFOAEs by Keefe et al. (2003, 2009). Kawase and Liberman (1993) and Micheyl and Collet (1996) also have considered the relationship between the efferent system and masking effects. Unlike the result reported here, Lilaonitkul and Guinan (2009a) did observe a rising, dynamic response using a bandpass noise centered on their tone, but the bandwidth of their noise was one-half octave, meaning that it encroached upon the frequency region characterized as lowpass here.

Strickland (2001,2004,2008) has provided rigorous, detailed analyses of the presumed mechanisms underlying the phenomenon of behavioral overshoot by supposing that the efferent system acts to alter the gain of the cochlear amplifier system, and thus to affect the

input/output function of the cochlea (also see Keefe et al., 2009). The basic assumptions are that cochlear output increases approximately linearly at low and high sound-pressure levels but increases more slowly over a middle range of sound-pressure levels. That is, the cochlear response is compressive over its middle range (Oxenham and Bacon, 2004). Furthermore, it is assumed that the magnitude of the compression (the gain of the cochlear-amplification system) is being continuously adjusted depending upon the recent history of stimulation. Specifically, after a period of relative silence, the gain is high, as is the degree of compression, and the onset of particular sounds leads to a reduction in the gain and a lessening of the compression over that middle range of sound-pressure levels. The adjustments in gain are presumed to be attributable to the activation of the MOC efferent system and the consequent inhibitory effect that such activation has on the electromotility of the outer hair cells (OHCs). In the case of overshoot, the onset of the wideband masking noise is assumed to lead to a reduction in gain that takes milliseconds to begin (the hesitation) and tens of milliseconds to complete (the dynamic response). Accordingly, the gain is still high during the presentation of tonal signals having short onset delays; thus, the signal-to-noise ratio is relatively low, and detectability is correspondingly poor. For long-delay signals, the gain (and compression) has had time to decrease considerably, the signal-to-noise ratio is higher, and detectability is better. In order for the nSFOAE responses we have reported here to conform with the above account, they would have to increase in magnitude for each decrease in gain and compression in the tone-plus-noise condition. This apparently paradoxical prediction of a larger nonlinear response in the face of decreasing cochlear nonlinearity is discussed further below.

A diagram of this Guinan/Strickland interpretation (borrowed from Strickland) is presented in Fig. 10. After a few hundred milliseconds in relative silence, the gain of the cochlear-amplifier system is assumed to be set high and the relevant input/output function is the highly compressive one at the far left (designated Tone-Alone). When only a single tone is turned on, there apparently is no activation of the MOC system, and thus no change in gain or in the amount of compression, so the initial input/output function is relevant for all three presentations of each triplet for tone-alone. Both of the single-earphone presentations of each triplet lead to the cochlear output associated with a 60-dB input (see horizontal dashed line labeled 1 in Fig. 10), and the final, two-earphone presentation of each triplet leads to the cochlear output associated with a 66-dB input (see dashed line labeled 2 in that figure), the latter of which falls considerably short of the sum of the first two outputs because of the strong compression (see expanded insert). To repeat, the discrepancy between the actual response observed and the response predicted by strict additivity is the nSFOAE measure reported here.

When the tone is accompanied by a noise of adequate level, the Guinan/Strickland explanation assumes that the MOC efferent system is activated, leading to a decrease in the gain of the cochlear amplifiers and an accompanying shift toward input/output functions that are less compressive than the initial function. The level of the noise determines which input/output function is most relevant for each presentation within the triplets. Because the two single-earphone presentations of each triplet are identical, the same point (see line labeled 1N in Fig. 10) on the same input/output function (designated Tone + Noise) is relevant for both single-earphone presentations. During the third presentation of each triplet, when both earphones are activated with exactly the same waveform, the sound in the ear canal is approximately 6 dB greater than during either of the two single-earphone presentations (see line labeled 2N in Fig. 10), meaning that the cochlear gain is reduced more than it is during those two other presentations. Thus, the relevant input/output function [designated (Tone + Noise) X 2 in Fig. 10] is slightly less compressive than the one relevant during those two single-earphone presentations, and the nSFOAE response is greater (see the expanded insert) than it would be if only the operating point were changed on the same input/output function

(lines labeled 1 and 2 illustrate one example). Note that the data shown in Fig. 5 reveal that it is the noise components below the frequency of the tone that are primarily responsible for activating the MOC system in the region centered on the frequency of the tone. To the extent that that band of frequencies is reasonably wide, the magnitude of the reduction in gain should show relatively little moment-to-moment fluctuation during the individual presentations of each triplet.

Note that this description provides a solution to what initially might appear to be an inconsistency between the data and the explanation offered. Namely, the explanation says that when the efferent system is activated, the cochlea moves toward a more linear, less compressive input/output function, yet the data show the magnitude of the nSFOAE increasing, not decreasing, when noise is presented. The expanded insert in Fig. 10 shows how both can be true. The nSFOAE increases because when tone-plus-noise is presented over both earphones simultaneously (line 2N), the relevant input/output function is different in two ways from when tone-plus-noise is presented only over one earphone (line 1N): the function is more linear, and it has lower overall gain. Thus, the difference that defines the nSFOAE is larger than when the same input/output function is relevant for all three presentations of each triplet. (Although the nSFOAE response necessarily reveals *only* the net nonlinear behavior of the SFOAE, Fig. 10 reveals that the magnitude of the nSFOAE obtained clearly can be *affected* by the processes, both linear and nonlinear, that determine which input/output functions are relevant.)

Although the above argument may account reasonably well for the data we have obtained, it appeals only to the amplitude compression that occurs in early cochlear processing. It is likely that phase (timing) effects also are at play (see Guinan, 2006). When the MOC system is activated, the local stiffness of the cochlear partition changes, meaning that each affected location is re-tuned to a slightly different frequency. Under certain assumptions, this shift in tuning would be accompanied by a change in round-trip propagation time. Indeed, when a moving-window analysis similar to that described above for amplitude was applied to the phase of the 4.0-kHz component of the nSFOAE response to tone-plus-noise, the phase function exhibited a temporal pattern similar to the rising, dynamic response for nSFOAE amplitude. Specifically, the phase typically increased about 40 – 100 degrees over the initial 150 ms of tone-plus-noise presentation. Also note that the existence of a phase difference between the two-earphone presentation and the single-earphone presentations allows the nSFOAE response to be larger than if no phase difference existed. Additional work on the effects of phase is ongoing.

Apparently, after the termination of a noise, the resetting of the gain and the transition back toward the initial state occur relatively slowly, so tones presented after the termination of a noise (Fig. 6) continue to be processed by those input-output functions active during the noise for an additional several hundred milliseconds. This long resetting time is supported by the report of Wiederhold and Kiang (1970) that the decay of efferent activity in primary auditory nerve fibers is much slower than is its onset. Also, McFadden (1989) reported that several hundred milliseconds of silence were required to reset the auditory system in an overshoot task. [For comparison, Backus and Guinan (2006, Table II) reported time constants for decay of their response that were only slightly longer than the short time constants fitted to the rise of their response.] All the data reported here were collected with knowledge of this slow resetting in mind; because of it, the offset of every noise burst was separated from the onset of the next noise burst by at least 500 ms for both the physiological and behavioral measurements.

The effects of the MOC efferent system can be observed in one ear even when the activating sound is presented only to the opposite ear (e.g., Guinan, 2006; Lilaonitkul and Guinan,

2009b; Walsh et al., 2010). For example, the same pattern of results shown in Fig. 5 also was obtained in all four subjects who were tested with the noise bands in the contralateral ear. If behavioral overshoot is a by-product of the functioning of the MOC efferent system, then one would expect overshoot also to be observable when a gated noise is presented to the contralateral ear. While studied only rarely, behavioral overshoot using contralateral noise bands has been reported, at least for some subjects (Turner and Doherty, 1997; Bacon and Liu, 2000).

D. The anomalous subject

One of the strongest arguments against the nSFOAE response being correlated with the processes underlying the behavioral phenomenon of overshoot is subject KW, whose physiological response was typical of the other subjects but whose behavioral overshoot was essentially zero with the wideband noise. Because a single instance of this sort can be argued to be fatal to any prospect of the nSFOAE response being indicative of the early stages of auditory processing for the overshoot task, we need to know as much as possible about this subject.

This absence of behavioral overshoot in subject KW using the wideband noise was documented in numerous test sessions distributed over more than a year, as was the presence of about 3.0 dB of overshoot using the lowpass noise. When additional behavioral data were collected for KW using the wideband noise and test tones of different frequency, overshoot magnitude was −0.2, −0.3, 2.9, 0.4, and 2.3 dB for signal frequencies of 1.8, 3.6, 3.9, 4.0, and 4.1 kHz, respectively. That is, overshoot was quite small for several frequencies in the vicinity of the 4.0-kHz tone used for the other subjects (we were looking for possible microstructure in overshoot). Past findings have linked small values of overshoot to temporary and permanent hearing loss (e.g., Champlin and McFadden, 1989; McFadden and Champlin, 1990; Bacon and Takahashi, 1992; Turner and Doherty, 1997; Strickland and Krishnan, 2005). However, among the four subjects, KW was second best at detection in the quiet at 4.0 kHz; on the other hand, he did have only one weak SOAE in the ear tested, which commonly is considered to be a sign of weak cochlear amplifiers. The time constant estimated for the dynamic nSFOAE response for subject KW was intermediate in size for these seven subjects (65 ms; see Fig. 4).

McFadden (1989) and Overson et al. (1996) showed that short silent periods between the offset of one noise burst and the onset of another (e.g., the silent period between the first and second observation intervals in psychophysical tests), can produce diminished magnitudes of overshoot during that second interval, presumably because the auditory system needs time to fully "reset" to its resting state. So, one possible explanation for why subject KW had little or no overshoot is that, for some reason, his auditory system needs longer to "reset" than those of the other subjects studied. However, the data in Fig. 6 suggest that KW's resetting time was not atypically long. Note that Keefe et al. (2009) also used long silent intervals between presentations, and future investigators are encouraged to do the same, both behaviorally and physiologically.

Because subject KW also was the primary experimenter, it is logically possible that his knowledge about the overshoot phenomenon and about the measurements being made somehow altered his auditory system in a way that allowed him to perform atypically well in the short-delay condition. Contradicting this idea is the fact that KW's physiological data were much like those of the other subjects, and he did exhibit a modest overshoot at 4.0 kHz with the lowpass noise (Fig. 8). In some recent work, subject KW did have about 4.0 dB of overshoot at 4.0 kHz when the wideband masking noise was strongly amplitude-modulated, and about 9.0 dB when the amplitude-modulated noise was lowpass.

Note that large individual differences in overshoot magnitude among nominally normal-hearing subjects have been reported regularly over the years (summarized by Overson et al., 1996). It may be that hearing sensitivity is not the best test for damage to whatever auditory mechanisms are responsible for overshoot.

E. Conclusions

The nSFOAE response exhibits considerable similarity to the behavioral phenomenon of overshoot in auditory masking. When tone plus noise are presented, the nSFOAE response has a rising, dynamic segment reminiscent of the improvement in detectability observed when a signal is increasingly delayed from the onset of a noise burst; the time constants of the two effects are both short; wideband and lowpass noises both produce a rising, dynamic response physiologically and considerable overshoot behaviorally; bandpass noise produces neither a rising, dynamic response physiologically nor overshoot behaviorally; there is some indication that stronger noise levels lead to smaller differences between the short- and long-delay conditions both behaviorally and in the nSFOAE (Walsh et al., 2010, Fig. 8); the nSFOAE requires a silent period of hundreds of milliseconds to recover after the offset of a noise, and behavioral overshoot requires hundreds of milliseconds of silence to reset; the rising, dynamic response and resetting both depend upon frequency components remote from the signal frequency; and the nSFOAE can be activated by a contralateral noise, and behavioral overshoot has been reported in some subjects tested with contralateral noise (Turner and Doherty, 1997; Bacon and Liu, 2000). These similarities, along with the effects of aspirin and noise exposure on overshoot, suggest that the normal functioning of the cochlea is at least partially responsible for the phenomenon of overshoot, and that the nSFOAE is somehow related to the information used by human listeners when detecting signals in the short- and long-delay conditions. Guinan (2006) and Kawase and Liberman (1993) also have concluded that cochlear mechanics and the MOC system can contribute to the detection of signals in noise, and that premise is at the heart of the work done by Strickland (2001, 2004, 2008).

On the other hand, there also were differences between the behavioral and nSFOAE measures. One difference was that the use of a highpass noise failed to produce a substantial dynamic nSFOAE response in any of the seven subjects (Fig. 5) even though that noise did produce large values of overshoot behaviorally for two of those subjects (Fig. 8). It is not at all difficult to accept that some of the defining characteristics of a complex behavioral phenomenon like overshoot could be determined at the cochlea, and some other defining characteristics could be determined beyond the cochlea, and this lack of similar effect of highpass noise on physiology and behavior in some subjects might be an example of this. After all, an entire nervous system does lie between the cochlea and behavior. If this is the right interpretation, then we have made progress on the task of determining some characteristics of overshoot that originate at early stages of processing and some that do not, and the focus should move to determining *where* along the auditory stream the effects of highpass noise are added to the characteristics bestowed by the cochlea.

A more concerning dissimilarity between physiology and psychophysics was the one subject who had essentially no overshoot behaviorally, but showed perfectly typical nSFOAE responses. In isolation, that outcome would suggest that the physiological measure under consideration is not truly tapping into the stream of processing that is relevant to performance in the overshoot task, raising doubt about the relevance of our nSFOAE measure for overshoot and other applications. However, adopting that interpretation would require our ignoring the numerous similarities between the physiological and psychophysical measures, and, for us at least, it is difficult to believe that all of those similarities are simply coincidences. At this time, we are unable to resolve the contradiction shown by this single subject. The best we can offer by way of a summary then is that some

characteristics of behavioral overshoot do appear to be obligatory consequences of human cochlear mechanics, but our nSFOAE measure is likely to be only a first step toward some future measure that will provide a clearer window on the cochlear contributions to behavioral phenomena like overshoot.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

List of Abbreviations

DPOAE	distortion-product otoacoustic emission
MER	middle-ear reflex
MOC	medial olivocochlear
OAE	otoacoustic emission
OHCs	outer hair cells
nSFOAE	nonlinear stimulus-frequency otoacoustic emission
SFOAE	stimulus-frequency otoacoustic emission
SOAE	spontaneous otoacoustic emission
TEOAE	transient-evoked otoacoustic emission

REFERENCES

Bacon SP. Effect of masker level on overshoot. J. Acoust. Soc. Am 1990;88:698–702. [PubMed: 2212293]

Bacon SP, Liu L. Effects of ipsilateral and contralateral precursors on overshoot. J. Acoust. Soc. Am 2000;108:1811–1818. [PubMed: 11051507]

Bacon SP, Takahashi GA. Overshoot in normal-hearing and hearing-impaired subjects. J. Acoust. Soc. Am 1992;91:2865–2871. [PubMed: 1629479]

Bacon SP, Repovsch-Duffey JL, Liu L. Effects of signal delay on auditory filter shapes derived from psychophysical tuning curves and notched-noise data obtained in simultaneous masking. J. Acoust. Soc. Am 2002;112:227–237. [PubMed: 12141348]

Backus BC, Guinan JJ Jr. Time-course of the human medial olivocochlear reflex. J. Acoust. Soc. Am 2006;119:2889–2904. [PubMed: 16708947]

Bassim MK, Miller RL, Buss E, Smith DW. Rapid adaptation of the 2f1-f2 DPOAE in humans: binaural and contralateral stimulation effects. Hear. Res 2003;182:140–152. [PubMed: 12948609]

Brownell WE, Bader CR, Bertrand D, de Ribaupierre Y. Evoked mechanical responses of isolated cochlear outer hair-cells. Science 1985;227:194–196. [PubMed: 3966153]

Carlyon RP. Changes in the masked thresholds of brief tones produced by prior bursts of noise. Hear. Res 1989;41:223–236. [PubMed: 2808151]

Champlin CA, McFadden D. Reductions in overshoot following intense sound exposures. J. Acoust. Soc. Am 1989;85:2005–2011. [PubMed: 2732381]

Church GT, Cudahy EA. The time course of the acoustic reflex. Ear Hear 1984;5:235–242. [PubMed: 6468781]

Dallmayr C. Stationary and dynamical properties of simultaneous evoked otoacoustic emissions (SEOAE). Acustica 1987;63:243–255.

Dallos, P. The auditory periphery: biophysics and physiology. Academic Press; New York: 1973.

Davis H. An active process in cochlear mechanics. Hear. Res 1983;9:79–90. [PubMed: 6826470]

Goodman SS, Keefe DH. Simultaneous measurement of noise-activated middle-ear muscle reflex and stimulus frequency otoacoustic emissions. J. Assoc. Res. Otolaryngol 2006;7:125–139. [PubMed: 16568366]

Goodman SS, Withnell RH, Shera CA. The origin of SFOAE microstructure in the guinea pig. Hear. Res 2003;183:7–17. [PubMed: 13679133]

Guinan JJ Jr. Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear. Hear 2006;27:589–607. [PubMed: 17086072]

Guinan JJ Jr. Backus BC, Lilaonitkul W, Aharonson V. Medial olivocochlear efferent reflex in humans: otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs. J. Assoc. Res. Otolaryngol 2003;4:521–540. [PubMed: 12799992]

Hicks ML, Bacon SP. Factors influencing temporal effects with notched-noise maskers. Hear. Res 1992;64:123–132. [PubMed: 1490895]

Jesteadt W, Bacon SP, Lehman JR. Forward masking as a function of frequency, masker level, and signal delay. J. Acoust. Soc. Am 1982;71:950–962. [PubMed: 7085983]

Kawase T, Liberman MC. Antimasking effects of the olivocochlear reflex. I. Enhancement of compound action potentials to masked tones. J. Neurophysiol 1993;70:2519–2532. [PubMed: 8120596]

Keefe DH. Double-evoked otoacoustic emissions. I. Measurement theory and nonlinear coherence. J. Acoust. Soc. Am 1998;103:3489–3498.

Keefe DH, Ellison JC, Fitzpatrick DF, Gorga MP. Two-tone suppression of stimulus frequency otoacoustic emissions. J. Acoust. Soc. Am 2008;123:1479–1494. [PubMed: 18345837]

Keefe DH, Schairer KS, Ellison JC, Fitzpatrick DF, Jesteadt W. Use of stimulus-frequency otoacoustic emissions to investigate efferent and cochlear contributions to temporal overshoot. J. Acoust. Soc. Am 2009;125:1595–1604. [PubMed: 19275317]

Keefe DH, Schairer KS, Jesteadt W. Is there an OAE correlate to behavioral overshoot? Abstr. Assoc. Res. Otolaryngol 2003;26:397.

Kemp DT. Stimulated acoustic emissions from within the human auditory system. J. Acoust. Soc. Am 1978;64:1386–1391. [PubMed: 744838]

Kemp DT. Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. Arch. Otol. Rhinol. Laryngol 1979;224:37–45.

Kemp DT. Towards a model for the origin of cochlear echoes. Hear. Res 1980;2:533–548. [PubMed: 7410259]

Kemp, DT.; Chum, RA. Observations on the generator mechanism of stimulus frequency acoustic emissions—Two tone suppression. In: van den Brink, G.; Bilsen, FA., editors. Psychophysical, Physiological, and Behavioral Studies in Hearing. Delft University; Delft, The Netherlands: 1980. p. 34-42.

Kim DO, Dorn PA, Neely ST, Gorga MP. Adaptation of distortion product otoacoustic emission in humans. J. Assoc. Res. Otolaryngol 2001;2:31–40. [PubMed: 11545148]

Levitt H. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am 1971;49:467–477. [PubMed: 5541744]

Lilaonitkul W, Guinan JJ Jr. Reflex control of the human inner ear: a half-octave offset in medial efferent feedback that is consistent with an efferent role in the control of masking. J. Neurophysiol 2009a;101:1394–1406. [PubMed: 19118109]

Lilaonitkul W, Guinan JJ Jr. Human medial olivocochlear reflex: effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths. J. Assoc. Res. Otolaryngol 2009b;10:459–470. [PubMed: 19263165]

Lonsbury-Martin BL, Harris FP, Stagner BB, Hawkins MD, Martin GK. Distortion-product emissions in humans: II. Relations to acoustic immittance and stimulus-frequency and spontaneous otoacoustic emissions in normally hearing subjects. Ann. Otol. Rhinol. Laryngol. Suppl 1990;147:15–29. [PubMed: 2110796]

Maison S, Durrant J, Gallineau C, Micheyl C, Collet L. Delay and temporal integration in medial olivocochlear bundle activation in humans. Ear Hear 2001;22:65–74. [PubMed: 11271977]

McFadden D. Spectral differences in the ability of temporal gaps to reset the mechanisms underlying overshoot. J. Acoust. Soc. Am 1989;85:254–261. [PubMed: 2921407]

McFadden D, Champlin CA. Reductions in overshoot during aspirin use. J. Acoust. Soc. Am 1990;87:2634–2642. [PubMed: 2373798]

McFadden D, Pasanen EG. Otoacoustic emissions and quinine sulfate. J. Acoust. Soc. Am 1994;95:3460–3474. [PubMed: 8046138]

Micheyl C, Collet L. Involvement of the olivocochlear bundle in the detection of tones in noise. J. Acoust. Soc. Am 1996;99:1604–1610. [PubMed: 8819856]

Nelson PC, Young ED. Enhancement of neural responses in the awake marmoset inferior colliculus to stimuli that induce perceptual enhancement. Abstr. Assoc. Res. Otolaryngol 2009;32:220.

Overson GJ, Bacon SP, Webb TM. The effect of level and relative frequency region on the recovery of overshoot. J. Acoust. Soc. Am 1996;99:1059–1065. [PubMed: 8609289]

Oxenham, AJ.; Bacon, SP. Psychophysical manifestations of compression: Normal-hearing listeners. In: Bacon, SP.; Fay, RR.; Popper, AN., editors. Compression: From Cochlea to Cochlear Implants. Springer; New York: 2004. p. 62-106.

Probst R, Lonsbury-Martin BL, Martin GK. A review of otoacoustic emissions. J. Acoust. Soc. Am 1991;89:2027–2067. [PubMed: 1860995]

Sachs MB, Kiang NYS. Two-tone inhibition in auditory-nerve fibers. J. Acoust. Soc. Am 1968;43:1120–1128. [PubMed: 5648103]

Schairer KS, Fitzpatrick D, Keefe DH. Input-output functions for stimulus-frequency otoacoustic emissions in normal-hearing adult ears. J. Acoust. Soc. Am 2003;114:944–966. [PubMed: 12942975]

Schairer KS, Keefe DH. Simultaneous recording of stimulus-frequency and distortion-product otoacoustic emission input-output functions in adult ears. J. Acoust. Soc. Am 2005;117:818–832. [PubMed: 15759702]

Schairer KS, Ellison JC, Fitzpatrick D, Keefe DH. Use of stimulus-frequency otoacoustic emission latency and level to investigate cochlear mechanics in human ears. J. Acoust. Soc. Am 2006;120:901–914. [PubMed: 16938978]

Schairer KS, Ellison JC, Fitzpatrick D, Keefe DH. Wideband ipsilateral measurements of middle-ear muscle reflex thresholds in children and adults. J. Acoust. Soc. Am 2007;121:3607–3616. [PubMed: 17552712]

Schmidt S, Zwicker E. The effect of masker spectral asymmetry on overshoot in simultaneous masking. J. Acoust. Soc. Am 1991;89:1324–1330. [PubMed: 2030219]

Shannon RV. Two-tone unmasking and suppression in a forward-masking situation. J. Acoust. Soc. Am 1976;59:1460–1470. [PubMed: 939879]

Shera CA, Guinan JJ Jr. Stimulus-frequency-emission group delay: a test of coherent reflection filtering and a window on cochlear tuning. J. Acoust. Soc. Am 2003;113:2762–2772. [PubMed: 12765394]

Smith DW, Moody DB, Stebbins WC, Norat MA. Effects of outer hair cell loss on the frequency selectivity of the patas monkey auditory system. Hear. Res 1987;29:125–138. [PubMed: 2442130]

Smith RL. Adaptation, saturation, and physiological masking in single auditory-nerve fibers. J. Acoust. Soc. Am 1979;65:166–178. [PubMed: 422812]

Smith RL, Zwislocki JJ. Short-term adaptation and incremental responses in single auditory-nerve fibers. Biol. Cybern 1975;17:169–182. [PubMed: 1125344]

Strickland EA. The relationship between frequency selectivity and overshoot. J. Acoust. Soc. Am 2001;109:2062–2073. [PubMed: 11386558]

Strickland EA. The temporal effect with notched-noise maskers: analysis in terms of input-output functions. J. Acoust. Soc. Am 2004;115:2234–2245. [PubMed: 15139634]

Strickland EA. The relationship between precursor level and the temporal effect. J. Acoust. Soc. Am 2008;123:946–954. [PubMed: 18247897]

Strickland EA, Krishnan LA. The temporal effect in listeners with mild to moderate cochlear hearing impairment. J. Acoust. Soc. Am 2005;118:3211–3217. [PubMed: 16334697]

Turner, CW.; Doherty, KA. Temporal masking and the "active process" in normal and hearing-impaired listeners. In: Jesteadt, W., editor. Modeling Sensorineural Hearing Loss. Erlbaum; Hillsdale, NJ: 1997. p. 387-396.

von Klitzing R, Kohlrausch A. Effect of masker level on overshoot in running- and frozen-noise maskers. J. Acoust. Soc. Am 1994;95:2192–2201. [PubMed: 8201115]

Walsh, KP. Unpublished Master's thesis. The University of Texas at Austin; 2009. Psychophysical and physiological measures of dynamic cochlear processing.

Walsh KP, Pasanen EG, McFadden D. Overshoot measured psychophysically and physiologically in the same ears. Abstr. Assoc. Res. Otolaryngol 2008;31:927.

Walsh KP, Pasanen EG, McFadden D. Evidence for dynamic cochlear processing in otoacoustic emissions and behavior. Abstr. Acoust. Soc. Am 2009;125(Pt. 2):2720.

Walsh KP, Pasanen EG, McFadden D. Properties of a nonlinear version of the stimulus-frequency otoacoustic emission. J. Acoust. Soc. Am. 2010 in press.

Westerman LA, Smith RL. Rapid and short-term adaptation in auditory-nerve responses. Hear. Res 1984;15:249–260. [PubMed: 6501113]

Whitehead ML. Slow variations of the amplitude and frequency of spontaneous otoacoustic emissions. Hear. Res 1991;53:269–280. [PubMed: 1880080]

Wiederhold ML, Kiang NYS. Effects of electric stimulation of the crossed olivocochlear bundle on single auditory-nerve fibers in the cat. J. Acoust. Soc. Am 1970;48:950–965. [PubMed: 5480390]

Wright BA. Detectability of simultaneously masked signals as a function of signal bandwidth for different signal delays. J. Acoust. Soc. Am 1995;98:2493–2503. [PubMed: 7593932]

Zwicker E. Temporal effects in simultaneous masking by white-noise bursts. J. Acoust. Soc. Am 1965a;37:653–656.

Zwicker E. Temporal effects in simultaneous masking and loudness. J. Acoust. Soc. Am 1965b; 38:132–141. [PubMed: 14347604]

Zwicker E, Schloth E. Interrelation of different oto-acoustic emissions. J. Acoust. Soc. Am 1984;75:1148–1154. [PubMed: 6725763]
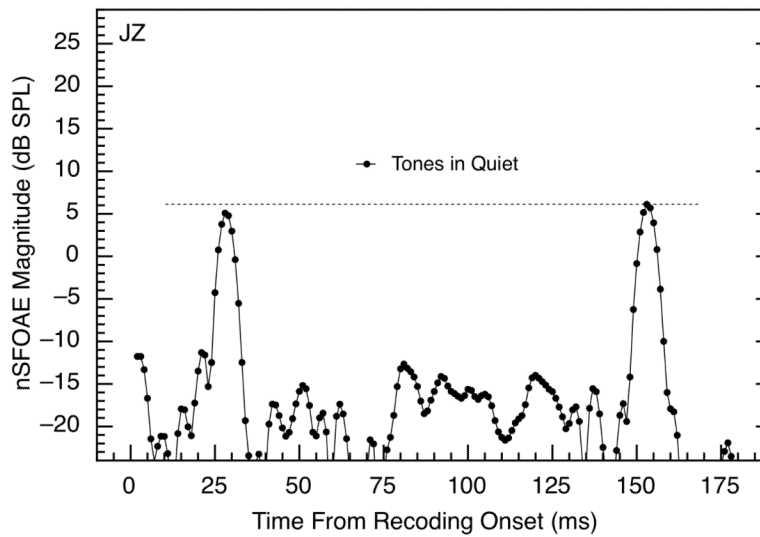
Fig. 1.
Magnitudes of nSFOAE responses to two brief 4.0-kHz tones of 60 dB SPL each and presented in the quiet. Onsets of the tones were at 25 and 150 ms from the onset of the recording period, and tone durations were 10 ms each (5 ms rise, 5 ms decay, with no steady-state segment). The dashed horizontal line reveals that the maximum magnitudes of the two nSFOAE responses were essentially identical. Data are based on 50 triplets obtained from subject JZ; the same outcome was obtained from other subjects.
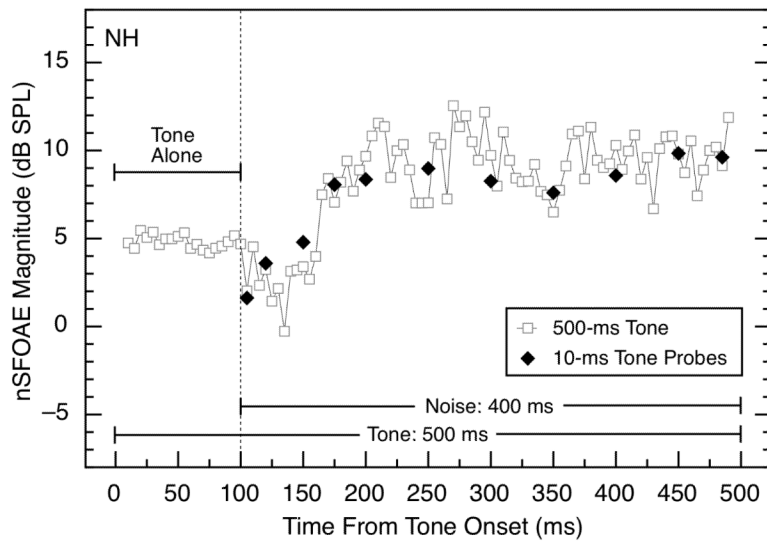
Fig 2.
Comparison of the magnitudes of nSFOAE responses to a long-duration (500-ms) tone or
short-duration (10-ms) tone bursts presented with differing time delays following onset of
the wideband noise. The tone was 4.0 kHz at 60 dB SPL, and the noise was 0.1 – 6.0 kHz
wide, 25 dB spectrum level, and 400 ms in duration. Results are shown only for subject NH,
but they are representative. For the 500-ms tone, points are plotted for 5-ms steps of the
analysis window. Some of the fluctuations in the nSFOAE response to the long-duration
tone are attributable to the envelope fluctuations in the specific noise sample used (see
Walsh et al., 2010, Figs. 2 and 3).
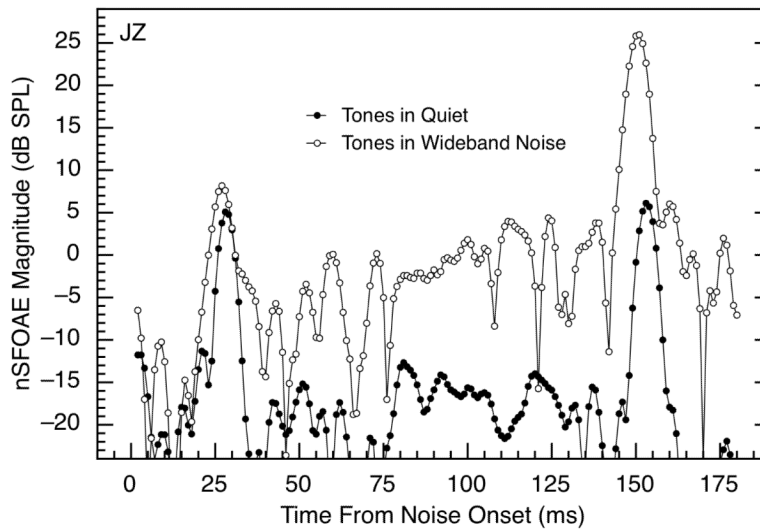
Fig. 3.
Magnitudes of nSFOAE responses to two brief tones presented in the quiet (solid symbols) or with a wideband noise (open symbols). Data for the tones in the quiet are replotted from Fig. 1. Results are shown for subject JZ, but they are representative. For purposes of this illustration only, the two tone bursts were separated by a time interval considerably shorter than the minimum used during actual data collection (which was 195 ms), and the remaining 200 ms of noise-alone following the second tone were omitted for clarity. The two sets of data shown here were obtained in the same test session.
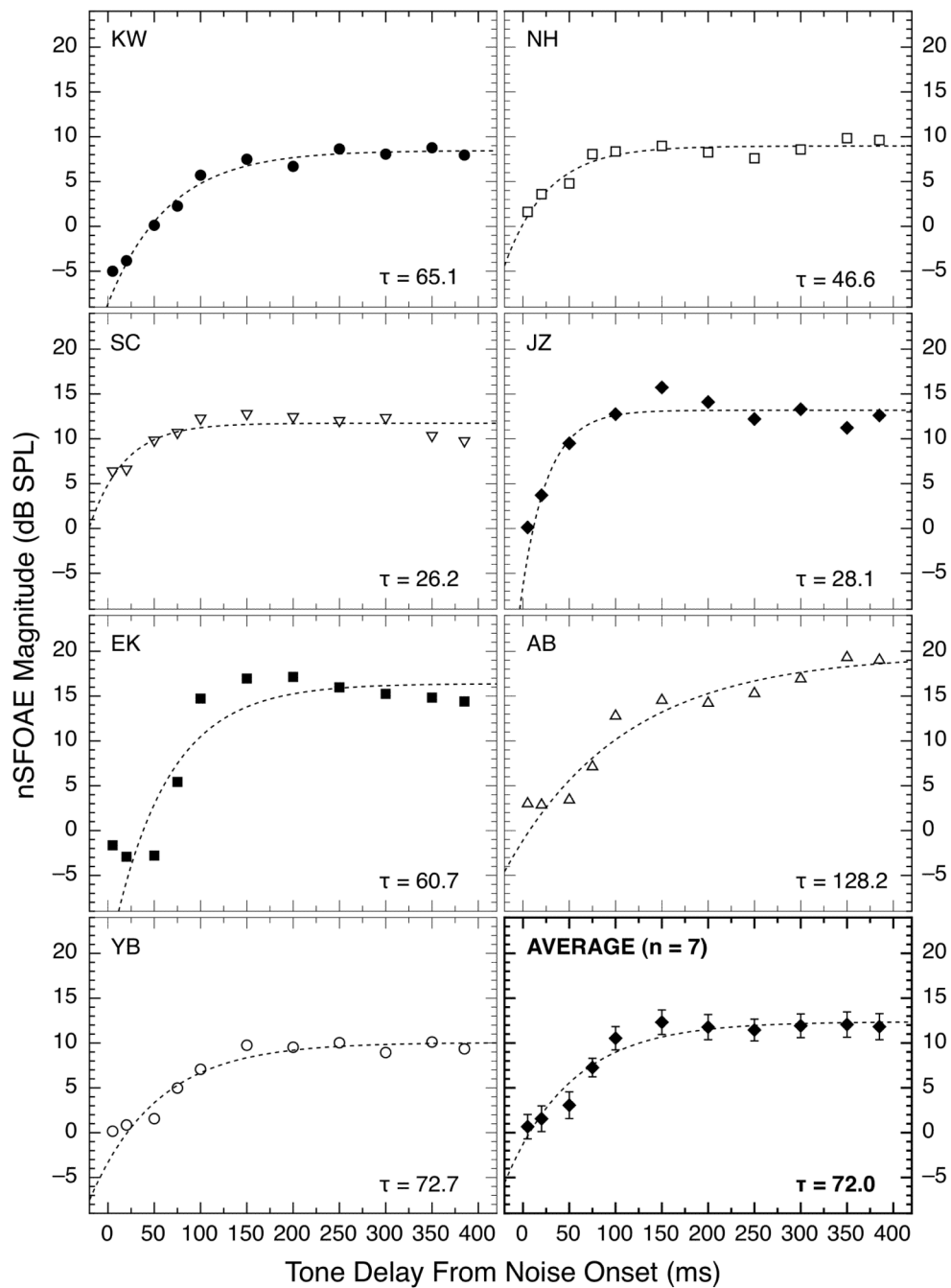
Fig. 4.
Gradually increasing magnitudes of the nSFOAE response to brief tones presented at differing times following the onset of the 400-ms wideband noise. For efficiency, these data were collected in blocks of trials in which every stimulus presentation contained two tone delays, always separated by at least 195 ms. The tone was 4.0 kHz, 60 dB SPL in level, and 10 ms in duration, and the noise was 0.1 – 6.0 kHz wide, 25 dB spectrum level, and 400 ms in duration. Each subject's data are shown in a separate panel; the four subjects at the top were the first crew of listeners tested. The dashed lines show the best-fitting positive exponential functions beginning 20 ms after noise onset (the hesitation). The weaker

nSFOAE response for short delays than for long delays is reminiscent of the behavioral phenomenon of overshoot.
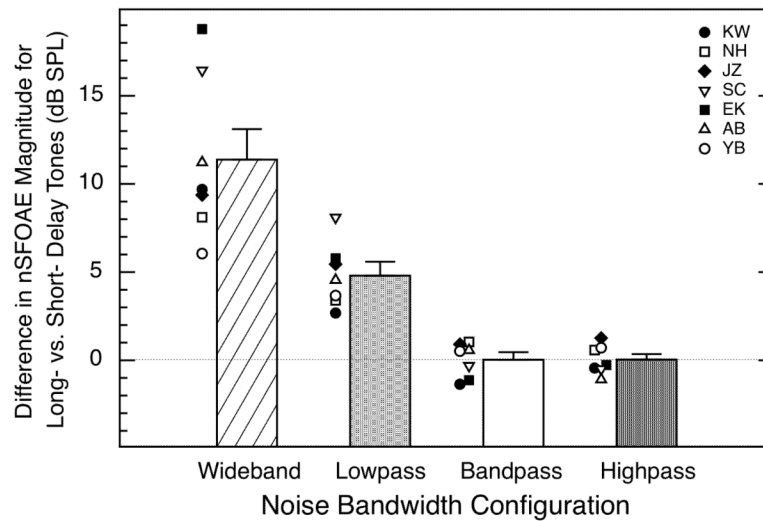
Fig. 5.
Differences in nSFOAE magnitudes for long-delay (200-ms) and short-delay (5-ms) tones plotted as a function of noise bandwidth. Bars indicate averages across all seven subjects, and symbols indicate individual data in each condition. The noise bandwidths were: 0.1 – 6.0 kHz (wideband), 0.1 – 3.8 kHz (lowpass), 3.8 – 4.2 kHz (bandpass), and 4.2 – 6.0 kHz (highpass); for all bandwidths, the spectrum level of the noise was 25 dB. Error bars indicate one standard error of the mean difference. For the highpass and bandpass noises, there was no rising, dynamic response to noise onset, and the responses to the short- and long-delay tones were essentially the same.
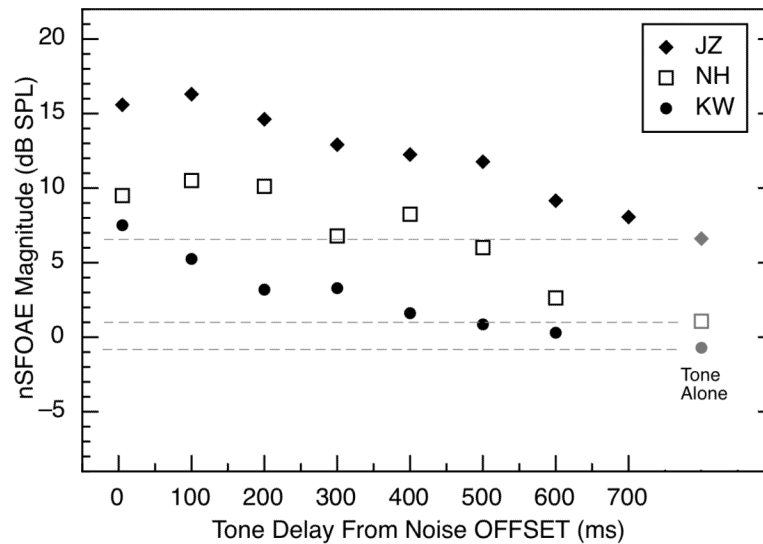
Fig. 6.
Magnitude of the nSFOAE response to brief tones presented at various time delays after the *offset* of a 200-ms wideband noise (25 dB spectrum level). Tones were 4.0 kHz, 60 dB SPL in level, and 10 ms in duration. The grey symbols at the right and the dashed horizontal lines show the magnitude of each subject's nSFOAE response to the 60-dB tone presented alone. Recovery to tone-alone levels was much slower than the rise to maximum magnitude shown in Fig. 4. For efficiency, these data were collected in blocks of trials in which every stimulus presentation contained two tone delays, always separated by at least 200 ms. The values plotted are the maximum magnitudes of the nSFOAE response to each tone presentation.
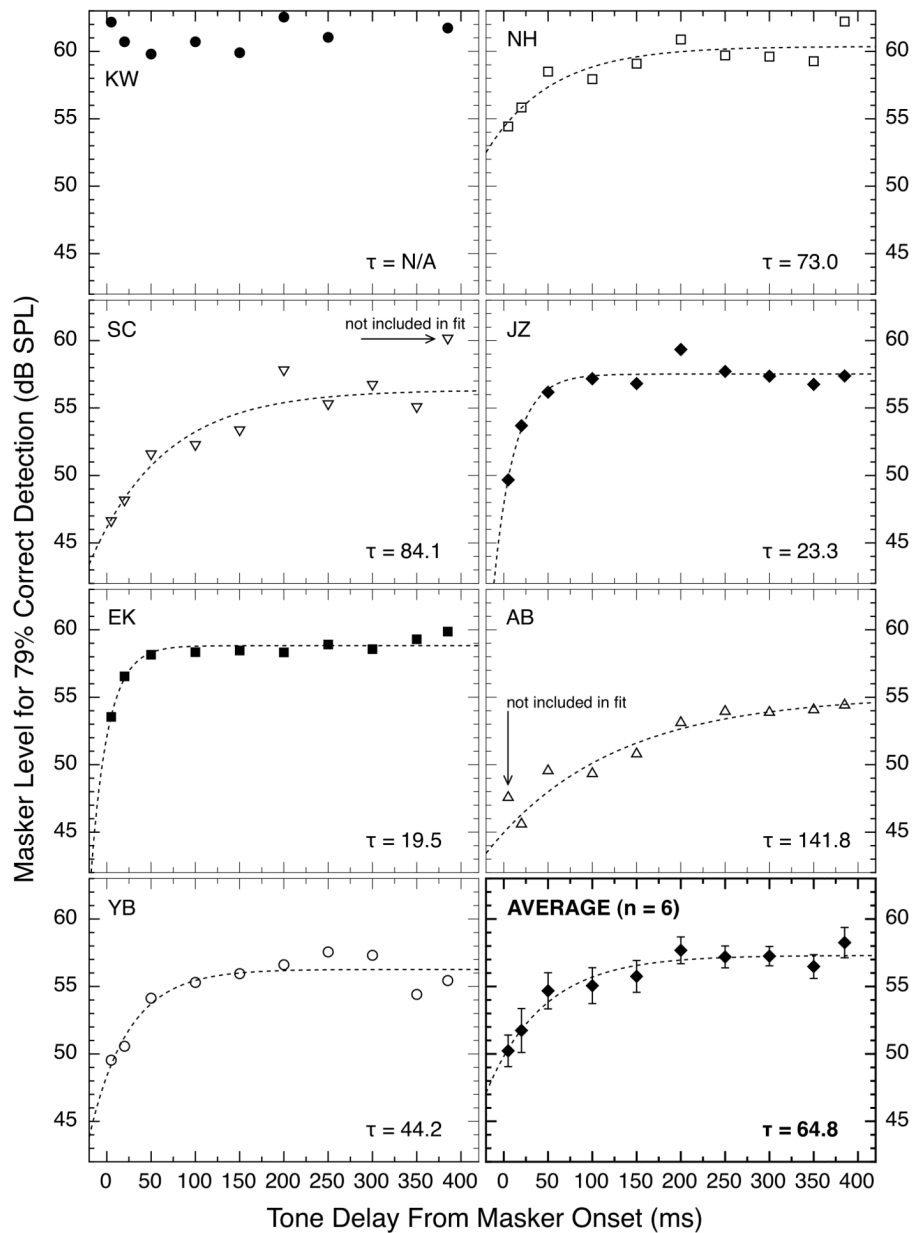
Fig. 7.
Masker levels required for 79% correct detections of a brief tonal signal are plotted as a function of signal delay from onset of the 400-ms masking noise. The signal was 4.0 kHz and 10 ms in duration; the noise was 0.1 – 6.0 kHz, 25 dB spectrum level, and 400 ms in duration. Each subject's data are shown in a separate panel. The dashed lines show the best-fitting positive exponential functions. For subject SC, the final data point was excluded when fitting the exponential function, and for subject AB the first data point was excluded from the fit (the fits were poor when these data were included).
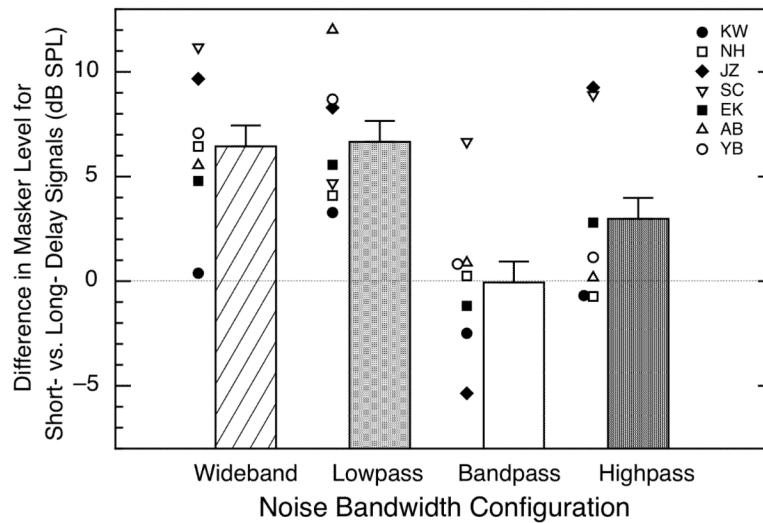
Fig. 8.
Differences in masker level necessary for a fixed level of detectability (79% correct) between short-delay (5-ms) and long-delay (200-ms) signals, using maskers of differing bandwidth. The signal was 4.0 kHz, 60 dB SPL, and 10 ms in duration. Bars indicate averages across all seven subjects and symbols indicate individual data in each condition. The noise bandwidths were: 0.1 – 6.0 kHz (wideband), 0.1 – 3.8 kHz (lowpass), 3.8 – 4.2 kHz (bandpass), and 4.2 – 6.0 kHz (highpass). Error bars indicate one standard error of the mean difference. For each bandwidth of the noise, at least 3 blocks of 50 trials each were collected for both the short- and long-delay conditions for every subject.
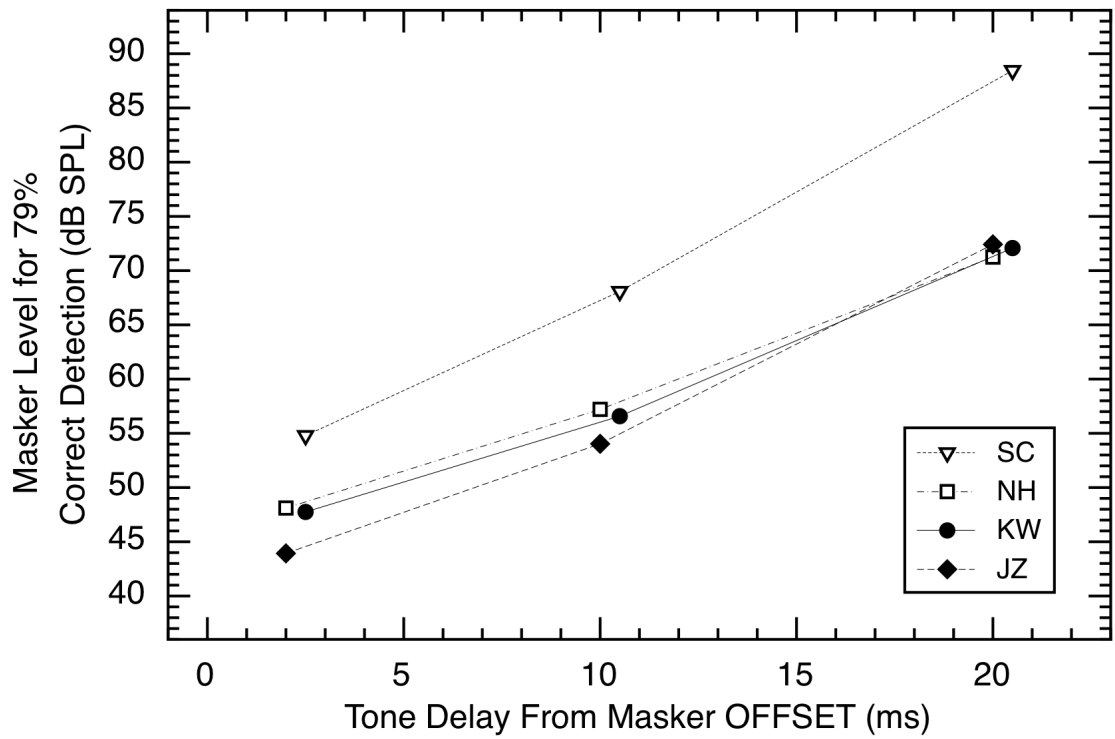
Fig. 9.
Masker levels required for 79% correct detections of a brief tonal signal presented at three temporal delays (2, 10, and 20 ms) after the *offset* of a wideband masker. The tone was 4.0 kHz, 45 dB SPL in level, and 10 ms in duration; the noise was wideband (0.1 – 6.0 kHz), 200 ms in duration, and varied adaptively in level. Some symbols have been laterally displaced slightly for clarity.
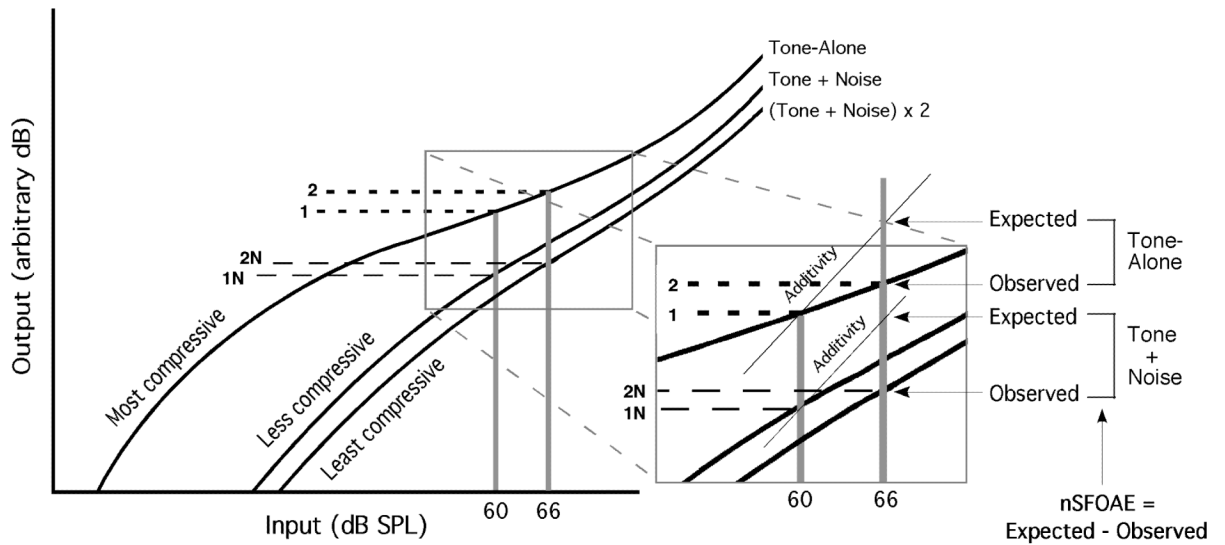
Fig. 10.
Diagram illustrating the input/output functions presumed to be relevant during stimulus conditions of the sort used here. When only tone-alone was presented, the same input/output function (designated Tone-Alone) was relevant for both the single-earphone and two-earphone presentations of each triplet; only the operating point on that function varied. When tone-plus-noise was presented, the cochlear gain was reduced and different input/output functions became relevant. The expanded insert shows the differences between Observed and Expected responses (which is the nSFOAE) for particular tone-alone and tone-plus-noise conditions.