

Published in final edited form as:

J Acoust Soc Am. 2012 October ; 132(4): 2524–2535. doi:10.1121/1.4751541.

Differences between psychoacoustic and frequency following response measures of distortion tone level and masking

Hedwig E. Gockel^{a)}, Redwan Farooq, and Louwai Muhammed

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, United Kingdom

Christopher J. Plack

Division of Human Communication and Deafness, The University of Manchester, Manchester M13 9PL, United Kingdom

Robert P. Carlyon

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, United Kingdom

Abstract

The scalp-recorded frequency following response (FFR) in humans was measured for a 244-Hz pure tone at a range of input levels and for complex tones containing harmonics 2-4 of a 300-Hz fundamental, but shifted by ± 56 Hz. The effective magnitude of the cubic difference tone (CDT) and the quadratic difference tone (QDT, at $F_2 - F_1$) in the FFR for the complex was estimated by comparing the magnitude spectrum of the FFR at the distortion product (DP) frequency with that for the pure tone. The effective DP levels in the FFR were higher than those commonly estimated in psychophysical experiments, indicating contributions to the DP in the FFR in addition to the audible propagated component. A low-frequency narrowband noise masker reduced the magnitude of FFR responses to the CDT but also to primary components over a wide range of frequencies. The results indicate that audible DPs may contribute very little to the DPs observed in the FFR and that using a narrowband noise for the purpose of masking audible DPs can have undesired effects on the FFR over a wide frequency range. The results are consistent with the notion that broadly tuned mechanisms central to the auditory nerve strongly influence the FFR.

I Introduction

In recent years, there has been increased interest in understanding the neural mechanisms underlying different aspects of auditory perception in humans. A potentially useful measure is the scalp-recorded frequency following response (FFR), which is an EEG recording that reflects sustained phase-locking of a population of neurons in the upper brain stem to stimulus-related periodicities (Marsh *et al.*, 1975; Smith *et al.*, 1975; Glaser *et al.*, 1976). It provides a non-invasive measure of neural processing in humans that can be compared to behavioral responses concerning the listener's perception. It has been argued that the FFR

^{a)}Author to whom correspondence should be addressed. hedwig.gockel@mrc-cbu.cam.ac.uk.

PACS number(s): 43.66.Hg, 43.66.Ki, 43.66.Dc [TD]

reflects processes important for the perception of pitch and that changes in the FFR with experience and/or training provide a measure of neural plasticity at the level of the brain stem (e.g., Greenberg *et al.*, 1987; Galbraith, 1994; Krishnan *et al.*, 2005; Russo *et al.*, 2008; Carcagno and Plack, 2011; but see Gockel *et al.*, 2011). The present study measures an important feature of the FFR, namely the presence of difference tones in response to multi-component complexes, and investigates the extent to which they might be related to the audible difference tones measured in psychoacoustic studies.

It is well known that difference tones, introduced by nonlinearities in the auditory system, can affect auditory perception. For example, when a harmonic complex consisting of harmonics above the fundamental frequency (F_0) is presented, the cubic difference tone (CDT; with frequency $2F_1 - F_2$, where F_1 and F_2 are the frequencies of two harmonics present in the complex and $F_2 > F_1$) can correspond to a harmonic at a frequency below those present in the signal. This can in turn increase pitch strength, although it is not essential for the perception of a pitch at the “missing F_0 ” (Houtsma and Goldstein, 1972). Similarly, the quadratic difference tone (QDT; $F_2 - F_1$) corresponds to a lower component, the largest being at the missing F_0 . Again, these components can affect pitch strength but are not essential for the ability to hear the missing- F_0 pitch (Licklider, 1956). A number of methods have been introduced to measure the CDT and QDT, including the introduction of an additional tone that either cancels, or beats with, the difference tone being investigated (e.g., Zwicker, 1955; Smoorenburg, 1972b; Buunen *et al.*, 1974; Pressnitzer and Patterson, 2001; Oxenham *et al.*, 2009). All of these methods assume that the difference tone is propagated from the site of generation in the cochlea to its characteristic place, where it acts “as if” it was present in the input signal.

Both the CDT and the QDT can also be observed in the frequency spectrum of the FFR to complex tones, and a number of studies using two-tone signals have characterized their dependence on stimulus parameters such as overall level and frequency separation (Rickman *et al.*, 1991; Krishnan, 1999; Pandya and Krishnan, 2004). For a harmonic complex, the frequency of the largest QDT is at the missing F_0 , and this component has been interpreted in terms of pitch perception (e.g., Greenberg *et al.*, 1987; Galbraith, 1994; Krishnan *et al.*, 2005; Russo *et al.*, 2008; but see Gockel *et al.*, 2011), as has the CDT in one study (Wile and Balaban, 2007). However, unlike the psychoacoustically measured difference tones, those observed in the FFR spectrum can arise from any stage of the auditory system where the responses to two or more components interact, provided that the neural response is related to the input signal by a nonlinear function (such as, for example, rectification or compression). Hence the extent to which combination tones in the FFR reflect the audible propagated combination tones that are measured in psychoacoustic studies is not obvious. In particular, inspection of the size of a combination tone in the FFR does not tell one whether it is consistent with the size of the propagated combination tones measured psychoacoustically. This is because perceptual levels are estimated in terms of the level of an *input signal*, expressed in SPL, whereas the difference tones measured in the FFR are an *electrophysiological response*, measured in microvolts. So, for example, a comparison of the level of a difference tone, relative to that of the primaries, in the two types of study will be hampered both by the often unknown shape of the function relating sound level to FFR amplitude, and by the lowpass frequency response of the FFR (e.g., Greenberg *et al.*, 1987).

We measured the FFR in response to a pure tone over a range of levels and determined the input level needed to produce a component in the FFR having the same level and similar frequency as the CDT and QDT generated by a three-component complex tone. This allowed us to measure the effective (input) level of the QDT and CDT in the same units as those used in psychoacoustic studies. The frequency components of the complex were derived by shifting harmonics two to four of a missing fundamental, allowing us to distinguish between the QDT and CDT (which, for a harmonic complex, would have had the same frequency). Three low harmonics were chosen to minimize peripheral interaction. We also examined the effect on the FFR of adding a narrowband noise as is typically used to mask audible DPs in both psychoacoustic and electrophysiological studies (Wile and Balaban, 2007; Carcagno and Plack, 2011; Krishnan and Plack, 2011). Our results revealed two important findings. First, the effective levels of both the QDT and CDT were considerably higher than estimates of the levels of the propagated difference tones obtained from psychoacoustic studies. Hence the QDT and CDT in the FFR reflect some additional interactions between, and nonlinear responses to, the components of the complex tone. These components were widely spaced and are usually assumed to be well resolved by the peripheral auditory system (Moore and Gockel, 2011). This suggests that these interactions either may have arisen in the low-frequency tails of peripheral neurons tuned to frequencies much higher than those of the primaries or may have occurred in broadly tuned neurons more centrally than the auditory periphery. Second, surprisingly, the lowpass noise not only reduced the FFR response to the difference tones but also the response to all three primary components of the complex, even though the upper cut-off of the noise spectrum was about 1.5 octaves below that of the highest component. We argue that this latter finding reflects the results of neural processes that operate over a wider frequency range than occurs at the level of the auditory nerve (AN).

II Methods

A Subjects

Six subjects (four female, two male) participated. They ranged in age from 19 to 23 yr and had self-reported normal hearing. They were selected from a pool of 10 subjects on the basis of initial FFR measurements for pure tones and complex tones (overlapping in only some cases partly with the stimuli used in the present study), where they were found to have robust FFR responses, i.e., clear peaks were observed in the magnitude spectrum of the FFR at the stimulus frequencies for moderate sound levels.

B Stimuli

The FFR was measured for two three-tone complexes. Both were derived from a harmonic complex tone containing harmonics two, three, and four of a 300-Hz F₀. In one, all harmonics were shifted downward by 56 Hz, i.e., the complex contained components at 544, 844, and 1144 Hz. In the other, all harmonics were shifted upward by 56 Hz, i.e., the complex contained components at 656, 956, and 1256 Hz. Thus the envelope rate (and QDT frequency) was identical (300 Hz) for the two complexes, but the CDT frequency, $2F_1 - F_2$, was 244 and 356 Hz for the downward- and upward-shifted complexes, respectively. The components had equal levels (70.2 dB SPL per component) and relative starting phases of

0°, 120°, and 240° for the bottom, middle, and top components, respectively. To estimate the effective level of the CDT of the downward-shifted complex, the FFR was measured for a 244-Hz sinusoidal tone presented at 60, 63, 66, 69, 72, and 75 dB SPL. All tones had a duration of 100 ms including 5-ms raised-cosine rise/fall times.

In one set of conditions, the tones were presented in quiet. The effect of a narrowband masking noise on the FFR was also investigated for the two complex tones and the 244-Hz tone at the 75-dB SPL level (see Table I). A digitally generated Gaussian noise in the frequency range 84–404 Hz (320 Hz wide, centered on 244 Hz) was presented with the 244-Hz sinusoidal tone. A second noise in the range from 140 to 440 Hz (300 Hz wide, centered on 290 Hz) was presented with the downward-shifted complex. A third noise in the range from 160 to 460 Hz (300 Hz wide, centered on 310 Hz) was presented with the upward-shifted complex. Noise segments 265.9-ms long with no onset or offset ramps were initially generated. They were played in loop mode, i.e., repeatedly without a silent gap, to give continuous noise. The frequency components in the noise were spaced by 3.761 Hz ($=1/0.2659$) to avoid audible clicks. The noise and the tones repeated in different non-integer time intervals (see below). Thus the starting phase of the noise differed across trials (while that of the tones was fixed), and the noise was effectively a “running” rather than a “frozen” noise. This allowed averaging of phase-coherent responses to the tones without averaging phase-coherent responses to the noise. For each condition, two independently sampled noises were used for the first and the second half of the trials. In the complex tone conditions, the noise (if present) had a root-mean-square (rms) level of 85 dB SPL. This level was chosen such that the stimuli in the complex-tone-with-noise conditions would be nearly identical to those used in a previous FFR study (Wile and Balaban, 2007). For the 244-Hz sinusoidal tone, the noise masker (if present) had rms levels of 86 and 90 dB SPL (spectrum levels of 61 and 65 dB re 20 μ Pa, respectively). The additional higher masker level was chosen because in psychophysical studies, the level of audible DPs is below the level of the primaries (see below). Thus the tone-to-masker ratio for the audible DPs in the complex-tone-with-noise conditions is lower than the tone-to-masker ratio for the sine tone with the 86 dB SPL noise. The 90 dB SPL noise level was chosen to reduce the tone-to-masker ratio in the pure tone condition, while avoiding uncomfortably loud levels.

Stimuli were generated with 16-bit resolution and a sampling rate of 40 kHz. They were played out through the digital-to-analog converter included in the evoked potentials acquisition system (Intelligent Hearing Systems-Smart-EP, IHS) and presented binaurally through mu-metal shielded Etymotic Research ER2 insert earphones, which have a flat frequency response at the human eardrum. In the conditions where the narrowband noise was present, the noise and the tone were defined as separate “channels” in the IHS system, before being mixed and sent to the earphones.

C Electrophysiological recording

Subjects rested comfortably in a reclining chair in a double-walled electrically shielded sound-attenuating booth. They were instructed to relax and to refrain from moving as much as possible during sound presentation and recording. They were allowed to fall asleep. The FFR was recorded differentially between gold-plated scalp electrodes positioned at the

midline of the forehead at the hairline (+, Fz) and the seventh cervical vertebra (–, C7). A third electrode placed on the mid-forehead (Fpz) served as the common ground. For this “vertical” electrode montage, the FFR is assumed to reflect sustained phase-locked neural activity from rostral generators in the brain stem (IC and LL, Marsh *et al.*, 1975; Smith *et al.*, 1975; Glaser *et al.*, 1976; Galbraith, 1994; Krishnan, 2006). Electrode impedances were less than 1 k Ω for all recordings. The FFR signal was recorded with a sampling period of 0.075 ms, amplified by a factor of 100 000 and bandpass filtered from 50 to 3000 Hz (6 dB/octave roll-off, resistor-capacitor filter). Epochs with voltage changes exceeding 31 μ V were automatically discarded and the trial repeated. The polarity of the tones, but not of the noise masker (when present) was alternated for each presentation, and alternate-polarity sweeps were recorded and averaged in separate data buffers by the SmartEP system. The tone stimuli were played with a repetition rate of 3.57/s, i.e., every 280.11 ms. The same stimulus was played in blocks of 2500 (valid) trials. Two blocks were run for each stimulus at different times during a session, and the FFR waveform was averaged across those two blocks to reduce possible effects of recording time within a session. Data were collected in three separate sessions for each subject. The three sessions were separated by 1.5 wk on average. Corresponding “with masker” and “no masker” conditions were run within the same session. The overall duration of a session, including electrode placement and breaks, was about 3 hr. Control recordings in which all of the same procedures were followed but with the tubes of the insert earphones blocked resulted in no signal above the noise floor at stimulus component, envelope, or distortion product frequencies in the subtraction waveform (see below) of the FFR.

D Analysis

Offline processing was done using MATLAB (The Mathworks, Natick MA). First, the averaged FFR responses for original-polarity and for inverted-polarity stimuli were either added or subtracted and the result divided by two, for each subject and condition. Addition of responses to alternating polarity stimuli enhances the representation of phase-locked activity to the envelope of the stimulus and minimizes the representation of phase locking to temporal fine structure. Subtraction of responses to alternating polarity stimuli enhances the representation of phase locking to temporal fine structure and minimizes the representation of phase-locked activity to the envelope (for a discussion, see Aiken and Picton, 2008). The resulting waveform was highpass and lowpass filtered at 150 and 2000 Hz (8th-order digital Butterworth filter; 3-dB down cutoff-frequencies), respectively. Further analysis was restricted to the time range from 12 to 100 ms after stimulus onset. The value of 12 ms was chosen such that the beginning of the FFR onset response was excluded (see below). For spectral analysis, the 88-ms waveform was zero-padded symmetrically to make up a 1-s signal, and the magnitude spectrum was calculated via a discrete Fourier transform. The magnitude spectrum is specified in decibels re 0.01 μ V. The peak height at a specific frequency of interest, e.g., at the frequency of the first primary in the signal, was measured as the highest magnitude present in the spectrum within a 12-Hz range centered on the expected component frequency. This range was the same as that used in a previous study (Gockel *et al.*, 2011) and was chosen to allow for the fact that a clear stimulus related FFR response above the noise baseline often gave a spectral peak at a frequency that differed by a few hertz from the stimulus frequency. Noise baseline was defined for each subject and

condition as the maximum spectral magnitude within the 12-Hz range centered on the frequency of interest, calculated for the 50 ms immediately before the tone was presented. Frequency domain averages were calculated for each condition. Averages across subjects' individual magnitude spectra were calculated rather than averages across subjects' FFR waveforms to avoid effects arising from possible differences in the onset delay of the FFR between subjects. Statistical analysis (analysis of variance, ANOVA, *t*-tests) was performed on the spectral magnitudes at the frequencies of interest (expressed in dB).

III Results

The latency of the unprocessed FFRs was about 9 ms for the complex tones and 10 ms for the 75-dB pure tone, estimated visually as the time point relative to stimulus onset of the first occurrence of a major amplitude excursion followed by a regular pattern in the FFR traces. This is in good agreement with the range of latencies reported in the literature for FFRs (Glaser *et al.*, 1976; Skoe and Kraus, 2010) and is consistent with a generation site at the level of the IC or LL.

Figure 1 shows the magnitude spectra of the FFRs, averaged across subjects. Conditions without and with noise are shown on the left- and right-hand sides, respectively. The solid and the dashed line indicate the spectra for the subtraction and the addition waveforms, respectively. For all conditions, the spectra of the subtraction waveforms show clear peaks at all of the frequencies of the components present in the stimulus. For the complex-tone stimuli, these are identified by the harmonic number followed by an arrow indicating the direction of the frequency shift that was imposed on the harmonics [Figs. 1(a) and 1(b), downward shifted by 56 Hz; Figs. 1(c) and 1(d), upward shifted by 56 Hz]. The spectra of the subtraction waveforms also show clear peaks at the CDT frequency, i.e., at 244 Hz for the downward-shifted complex tone and at 356 Hz for the upward-shifted complex tone. As expected, the magnitudes of the peaks at the CDT frequencies and at the pure-tone frequency [Figs. 1(e) and 1(f)] were clearly reduced in the presence of the masker (see below). For the downward-shifted complex tone [Fig. 1(b)], adding the masker reduced the CDT peak to below noise baseline for two subjects.

The spectra of the addition waveforms (for the complex tone conditions only) show clear peaks at 300 Hz, i.e., at the envelope rate and frequency of the QDT (and integer multiples of it), for both the upward- and the downward-shifted complex. In agreement with earlier reports (e.g. Rickman *et al.*, 1991), this QDT component in the addition waveform was larger than the CDT component in the subtraction waveform for all conditions. This was confirmed by the results of a three-way repeated-measures ANOVA,¹ with factors: Type of distortion product, direction of frequency shift and presence of noise. There was a significant effect of type of distortion product [$F(1,5) = 77.0, P < 0.001$]; the QDT component was on average 9.5 dB higher in level than the CDT component. Paired *t*-tests showed that the QDT magnitude was significantly larger than the CDT magnitude for all four complex tone

¹Throughout the paper, if appropriate, the Huynh-Feldt correction was applied to the degrees of freedom (Howell, 1997). In such cases, the original degrees of freedom and the corrected significance value are reported.

conditions [all $P < 0.01$; Bonferroni corrected and one-tailed, as based on previous reports, the QDT was expected to be larger than the CDT].

The QDT was also larger than the CDT when the magnitude spectra were first calculated separately for each onset polarity and then averaged across polarities. An ANOVA based on the spectra averaged across polarities showed a significant effect of type of distortion product [$F(1,5) = 98.2$, $P < 0.001$]; the QDT was on average 8.9 dB higher in level than the CDT. The results of paired t -tests based on the spectra averaged across polarities were the same as those based on the spectra of the addition and subtraction waveforms reported above. Indeed, and in agreement with the results of Rickman *et al.* (1991), the peak heights observed at the QDT and CDT frequencies in the spectra averaged across polarities differed little from those observed at the QDT frequency in the spectrum of the addition waveform (which was divided by two, see Sec. II) and at the CDT frequency in the spectrum of the subtraction waveform (divided by two), respectively (see Fig. 2).

We now consider the effects of the masking noise. Figure 3 shows the magnitude of the peak at 244 Hz in the spectrum of the FFR subtraction waveform for the 75-dB SPL pure-tone conditions with and without the noise masker. A one-way repeated-measures ANOVA showed a significant effect of the noise [$F(2,10) = 28.94$, $P < 0.001$]. The peak height was reduced by 7.2 and 9.2 dB in the presence of the 86- and the 90-dB SPL noises (paired t -tests: $P < 0.01$ and $P < 0.001$; one-tailed, as noise was expected to reduce phase locking to the 244-Hz tone), respectively. The peak height was significantly lower in the presence of the 90-dB than in the presence of the 86-dB SPL noise (paired t -test: $P < 0.025$; one-tailed, as more masking was expected for the higher noise level).

To assess the effects of the narrowband masker on the FFR of the complex tones, a three-way repeated-measures ANOVA (with factors component frequency, direction of frequency shift, and presence of noise) was calculated on the magnitudes at the frequencies of the CDT and the three primaries measured in the subtraction waveform and at the frequency of the QDT measured in the addition waveform. The results showed a significant effect of noise [$F(1,5) = 38.6$, $P < 0.01$]. The main effects of component frequency and direction of frequency shift were also significant [$F(1,5) = 26.9$, $P < 0.001$ and $F(1,5) = 10.8$, $P < 0.05$, respectively]. The peak heights were somewhat larger for the downward- than for the upward-shifted complex. The interaction between component frequency and direction of frequency shift was significant [$F(4,20) = 8.1$, $P < 0.01$], as was the interaction between component frequency and noise [$F(4,20) = 3.0$, $P < 0.05$]. The latter indicates that the effect of the masking noise depended on the component under consideration. However, neither the interaction between noise and direction of frequency shift [$F(1,5) = 0.02$, $P = 0.89$] nor the three-way interaction [$F(4,20) = 0.85$, $P = 0.44$] was significant. Because no interaction involving both factors of noise and direction of frequency shift was significant, in the following, the effect of the masking noise for the various component frequencies was averaged across the data for the downward- and the upward-shifted complex.

Figure 4 shows the reduction in magnitude caused by the noise at the frequencies of the QDT, the CDT, and the three primaries. The noise reduced the peak height most at the CDT, as expected, but the noise had an effect for all component frequencies, even for the third

primary (the shifted fourth harmonic), which had the largest frequency separation from the noise band. A one-way repeated-measures ANOVA with the size of peak reduction at the five frequencies (QDT, CDT, and three primaries) as input showed a significant main effect of frequency [$F(4,20) = 3.0$, $P < 0.05$]. A simple contrast with the CDT component as reference showed that the reduction at the CDT was larger than those for the first and the second primaries [$F(1,5) = 8.9$ and 8.3 , respectively; $P < 0.05$ for both], but it was not significantly different from that for the third primary and the QDT [$F(1,5) = 5.0$ and 4.5 , respectively]. In addition, paired t -tests showed that at each of the three primary frequencies, the magnitude was larger without the noise than in the presence of the noise (all $P < 0.01$, one tailed, as reduction due to noise was expected). Thus, surprisingly, the narrowband masker affected the FFR over a wide range of frequencies.

Consider next the FFRs for the conditions in which the 244-Hz sinusoid was presented at various levels (in silence) to allow estimation of the effective level of the CDT, which, for the downward-shifted complex, had a frequency of 244 Hz. Figure 5 shows, for each subject, the magnitude of the peak at 244 Hz in the spectrum of the subtraction waveform as a function of level. Cases in which the peak value for the time interval when the tone was presented did not exceed the noise baseline are indicated by a downward-pointing arrow next to the corresponding (empty) symbol at the baseline value. For the 60-dB SPL tone level, the peak height was not above the baseline for five of the six subjects, indicating that FFR threshold was not reached. For one subject, even the 63 and the 66 dB SPL tones were below FFR threshold. The solid line indicates the mean peak height, averaged across the data for those subjects whose data were above the baseline. Note that the functions are approximately linear for higher levels (slope of about one) but flatten off at lower levels with clear individual differences between the levels where it starts to flatten off.

To estimate the effective level of the CDT for the downward-shifted complex, the level of the pure tone was determined that gave the same value in the magnitude spectrum of the subtraction FFR at 244 Hz as that observed in the magnitude spectrum of the subtraction FFR for the downward-shifted complex; this was done by linear interpolation from each subject's FFR growth function for the pure tone. The individual and mean effective levels are given in Table II. The mean effective level was 65.5 dB SPL, which is only 4.7 dB below the level of the primaries in the complex. For two of the subjects (subjects 4 and 6), the effective level was only 1.5 dB below the level of the primaries. Thus the effective level of the CDT observed in the FFR is rather high relative to the input sound level of the primaries.

An estimate of the effective level of the QDT can also be obtained using the input-output curve measured for the 244-Hz sinusoid, based on the assumption that the 56-Hz difference between 244 Hz and the 300-Hz frequency of the QDT would not have a large effect on the FFR magnitude. Therefore we derived a rough estimate of the effective QDT level by finding the level of a 244-Hz sinusoid that, for each subject, produced the same FFR peak height as the average value observed at the QDT frequency in the magnitude spectra of the addition waveforms for the downward- and the upward-shifted complex tones (see Table III). For four of the six subjects, the effective level of the QDT was above 75 dB SPL, i.e., above the maximum level of the sinusoid tested. When the spectral magnitudes of the FFRs at the QDT frequency were first averaged across subjects (Table III, right-hand column, last

line) and then the level of a 244-Hz sinusoid that produced the same peak height (averaged across subjects, solid line in Fig. 5) was found, the estimated effective level of the QDT was also above 75 dB SPL. Thus, on average, the effective level of the QDT in the FFR was more than 5 dB above the level of the individual primaries in the complex (and more than 9.5 dB above the effective level of the CDT).

IV Discussion

The FFR was measured for two complex tones containing frequency-shifted harmonics two to four of a 300-Hz F0. The results showed: (i) In agreement with previous findings, the spectral amplitude of the FFR at the CDT frequency was smaller than at the QDT frequency, which is the opposite of what is typically found psychophysically. (ii) When a narrowband noise masker centered on the region of the DPs was added, the magnitude of the spectral component of the FFR at the CDT frequency was reduced, as expected. In addition, however, the magnitude of the QDT and that of all primary frequencies was significantly reduced by the noise, indicating that phase locking to the peripherally resolved primaries was affected. (iii) The input level of a sine tone at the CDT frequency that was needed to produce the same magnitudes as the spectral components at the CDT and QDT frequencies in the FFR (the effective level of the DPs) was 4.7 dB below and 5 dB above the level of the individual primaries, respectively. We next discuss the effective DP levels in the FFR relative to those reported in psychophysical experiments and then the effect of the narrowband noise masker.

A Comparison of effective DP levels in the FFR with psychoacoustic results

There have been many studies estimating the level of audible DPs. Here we focus on two studies whose conditions were most similar to those used here to estimate an upper limit of the level of the audible DPs for the present tone complexes.

1 CDT—Measurements of the audible CDT level for a complex tone with three harmonics were reported by Oxenham et al. (2009). They used the method of best beats (see below) to estimate the relative CDT level for a complex tone with a 222-Hz F0. Harmonics seven to nine were added in cosine phase and presented at 65 dB SPL. The CDT level was estimated to be about 14 dB below that of the level of the individual primaries. This is almost certainly higher than the audible CDT level for the tone complex used in the present study, for two reasons. (i) The frequency ratio of the two lowest harmonics was 1.14 for Oxenham *et al.* (2009) and 1.5 here; psychoacoustic estimates of CDT level decrease with increasing frequency ratio (Goldstein, 1967; Hall, 1972; Smoorenburg, 1972b; Zwicker, 1979, 1981). In fact, for the present frequency ratio of 1.5, the CDT would be at/ or just outside the limit of the audibility region (Goldstein, 1967; Smoorenburg, 1972a), and thus, difficult to measure. (ii) The method used by Oxenham *et al.* (2009) to estimate CDT level involved the simultaneous presentation of an additional tone with the complex; the complex may suppress the additional tone, leading to an overestimate of DP level (Smoorenburg, 1972b).

All of this means that the audible CDT level for the complex used in the present study can safely be estimated to be well below 56 dB SPL. Therefore the CDT level measured in the FFR for the present complex tone is at least 9 dB higher than the level of the audible CDT as

measured in psychoacoustic experiments. It is also worth noting that with the current subjects and measurement procedure, a 244-Hz sinusoid presented at 56 dB SPL would not lead to an FFR above baseline (see Fig. 5).

2 QDT—The effective level of the QDT at 300 Hz in the FFR was 5 dB above the input level of the primaries. In psychoacoustic studies, the level of the audible QDT is usually smaller than that of the CDT. Oxenham *et al.* (2009) estimated the audible QDT level for the 222-Hz F₀, 65-dB SPL complex with harmonics seven to nine added in cosine phase to be about 35 dB below that of the level of the individual primaries. Hall (1972) reported that for a two-tone complex with F₁ = 583 Hz, F₂ = 875 Hz, and primary tone levels of 68 dB SPL, a stimulus similar to that used here (apart from the number of components), the relative QDT level was about -33 dB. Pressnitzer and Patterson (2001) estimated the QDT level for a 100-Hz F₀ complex, with the lowest harmonic at 1500 Hz, as a function of the number of harmonics present in the complex tone. They reported an increase in the QDT level of about 2 dB when the number of harmonics present in the complex tone was increased from two to three. Allowing for the increase in QDT level with number of harmonics, therefore, the QDT level for a three-tone complex otherwise like the complex used by Hall (1972) would be at about -31 dB.

All of these studies show that for the complex tone used in the present experiment the level of the audible QDT must be at least 20 dB below the level of the primaries. Therefore the effective QDT level measured here in the FFR is at least 25 dB higher than the level of the audible QDT as measured in psychoacoustic experiments.

3 Contributions to DPs in the FFR—The above comparisons suggest that the magnitude spectrum of the FFR at the CDT and QDT frequencies mainly reflects contributions in addition to those giving rise to the audible DPs. Furthermore, these additional contributions must be large, leading to increases of at least 9 and 25 dB in the FFR DPs over the audible DP levels for the CDT and QDT, respectively.

Nonlinear processes in the inner ear, more specifically processes responsible for cochlear amplification, are thought to underlie the generation of the audible CDT and part of the QDT. For the QDT, there is probably an additional component generated in the middle ear (e.g., Hall, 1972; Zwicker and Martner, 1990). The cochlear mechanical origin of the DPs is assumed to be in the region where the primaries interact, close to the F₂ place (e.g. Furst *et al.*, 1988). This has sometimes been called the “generative” component of the DP (Wile and Balaban, 2007). The generative component causes a traveling wave on the basilar membrane (BM), propagating to the characteristic place of the DP and stimulating hair cells tuned to the DP frequency. This has been called the “propagated” component (Wile and Balaban, 2007). While the cochlear mechanical origin of the DP is assumed to be in the region where the primaries interact, the *audible* DP is assumed to result from the mechanical wave propagation to its characteristic place (e.g., Furst *et al.*, 1988). In other words, the perception of the DP is believed to result from a cochlear response at the characteristic place of the DP, and perceptual cancellation of the DP is presumed to reflect a reduction of this response to below absolute or masked threshold (Furst *et al.*, 1988; Hartmann, 1997).

Our results show that the effective levels of the CDT and QDT as measured in the FFR are much higher than those estimated in psychophysical studies and thus are likely to reflect additional contributions. These additional “non-propagated” components could include neural activity arising from the nonlinear BM response in the region where the primaries interact. Nonlinearities at any stage in the auditory pathway—BM vibration, inner hair cell transduction, and neural responses up to the site of FFR generation—may add to the FFR DPs whenever those stages respond to more than one frequency component; in response to a sinusoidal input a static nonlinearity would produce DPs only above the input frequency. Consistent with this, the FFR for a pure tone shows DPs only at higher harmonics. DPs arising through half-wave rectification are likely to contribute to the QDT, but not to the CDT, observed in the FFR. This is because the CDT in the FFR is determined from the FFR subtraction waveform to alternating polarity stimuli; discharges during the condensation phase of the inverted stimulus alternate with the discharges occurring during the condensation phase of the non-inverted stimulus and subtraction of the two discharge pattern will approximately “recreate” the stimulus waveform (Aiken and Picton, 2008). However, the peak height at the CDT frequency calculated from averaging spectra determined separately for the FFR waveforms measured for the two starting polarities, in which DPs due to half-wave rectification would not be cancelled, is similar to the peak height at CDT in the spectrum of the subtraction waveform (see Fig. 2, bottom), indicating that half-wave rectification makes only a small contribution to the measured CDT.

In the FFR literature, the origin of the CDT has been assumed to be a mechanical nonlinearity in the cochlea (Bhagat and Champlin, 2004). For the QDT, a different additional source has been considered (Bhagat and Champlin, 2004) in nonlinear neural mechanisms (not specified in detail) that extract the envelope of the stimuli (Chertoff *et al.*, 1992; Dolphin *et al.*, 1994; Arnold and Burkard, 1998). At the level of single units in the mammalian auditory system, (subcortical) phase locking to envelope fluctuations has been observed in the AN, the cochlear nucleus, the superior olivary complex, and the IC (for a review, see Frisina, 2001). Kuwada *et al.* (2002) noted that “at every level of the auditory system, neurons can temporally follow the envelopes of modulated signals” and that the highest envelope rate that neurons can follow decreases at higher levels in the auditory pathway. Thus, for complex tones containing unresolved harmonics, including amplitude-modulated tones, phase locking of neurons to the envelope is a likely contributor to the large QDT component observed in the FFR. However, for low harmonic numbers, like the ones used here, the situation is less clear. This is because these components are generally assumed to be resolved in the auditory periphery, leading to little interaction except for relatively weak responses from neurons tuned to frequencies centered between the components. The large FFR responses at the DP frequencies must be based on the responses of neurons that are responding to more than one stimulus component. This could happen either via the low-frequency tails of neurons tuned to frequencies much higher than the primaries, which have been suggested to be the main origin of the FFR (Dau, 2003), or via broadly tuned neurons with CFs close to the frequencies of primary components. The latter have been observed at various levels above the AN, but not in the AN itself (Young *et al.*, 1992; Wiegrebe and Winter, 2001). They are likely to be the consequence of monaural processes as they were not

observed in the FFR for dichotically presented stimuli containing only low numbered harmonics (Gockel *et al.*, 2011).

B Effect of narrowband noise on the FFR

1 Pure tone FFR—We next discuss the effect of the narrowband noise on the FFR of the 244-Hz sinusoid and its relation to the psychoacoustic masked threshold of the sinusoid. The present results largely agree with those of previous studies that measured the FFR in response to a sinusoid as a function of the level of an on-frequency noise and also measured the psychophysical masked threshold in the same listeners (Marsh *et al.*, 1975; Glaser *et al.*, 1976). Marsh *et al.* (1975) reported a complete disappearance of the FFR response when the noise level was 5–10 dB above the psychophysically determined masking threshold. The latter was determined using the methods of limits, both ascending and descending, and a 50% correct response criterion was adopted. The size of the FFR response was determined from the FFR waveform with a hand-operated odometer, a measure that is sometimes called “string length.” Glaser *et al.* (1976) determined the size of the FFR response from the averaged peak-to-peak value of the filtered FFR waveform. They reported that no detectable FFR was observed when the subject reported complete masking of the tone and that “when the intensity of the noise was lowered in 10 dB increments, the FFR became detectable near the level at which the subject reported just being able to hear the tone in the noise.”

In the present study, for the 90-dB SPL noise, the tone-evoked FFR was still visible [see Fig. 1(f)] and the 244-Hz spectral component was above the noise baseline for all but one subject. In this condition, the tone-to-noise ratio (TNR) in the equivalent rectangular bandwidth (ERB_N , a measure used as an estimate of the bandwidth of the auditory filter) centered on 244 Hz (ERB_N width of 51 Hz) was -7 dB (Moore *et al.*, 1997). This is 3 dB lower than the signal-to-noise ratio at the output of the auditory filter required for threshold at high masker levels, i.e., masked threshold measured in psychophysical experiments (Moore *et al.*, 1997). However, psychoacoustic threshold measurements usually adopt a criterion level around 75%, and thus the signal is *not* completely inaudible at threshold. In addition, a perfect match between FFR threshold and psychoacoustic threshold should not be expected as the visibility of a peak at the signal frequency in the spectrum of the FFR depends on the number of trials over which the signal-plus-noise waveform is averaged.

In summary, the results of FFR experiments, which for practical reasons are limited to a few thousand trials, largely agree in indicating that FFR thresholds for masked tones are either roughly equal to or somewhat below psychoacoustically measured masked thresholds. Increasing the number of trials would be expected to decrease FFR thresholds for masked tones as long as the noise still allows some phase locking to the tone, even if it would not lead to detection of the tone in a psychoacoustic experiment.

2 Complex tone FFR—We next discuss the effect of the narrowband noise on the DP and primary components in the spectrum of the FFR for the complex tones. As expected, the peak height was reduced at the CDT frequency. A comparison between the noise effects on the CDT and the noise effects on the 244-Hz 75-dB SPL sinusoid is problematic because of

the difference in SNRs; in the absence of the noise, the size of the CDT component in the FFR for the complex was smaller than that for the sinusoid.

The important and unexpected finding was that the noise reduced the peak magnitude in the spectrum of the FFR not only at the CDT frequency but also at the QDT frequency and at *all* primary frequencies; for the latter, the average size of the reduction was between 3.3 and 4.9 dB. Figure 6 shows excitation patterns for the stimuli in the complex-tone conditions, calculated using the model described by Moore *et al.* (1997). The auditory filters in this model are based on data obtained in simultaneous masking experiments using notched-noise and rippled-noise maskers (Patterson, 1976; Houtgast, 1977; Glasberg *et al.*, 1984). The excitation patterns are thought to resemble the internal representation of the stimuli. They show that the noise would produce partial masking of the lower primary, more so for the downward-shifted complex [Fig. 6(a), top] than for the upward-shifted complex [Fig. 6(b), bottom]. However, for auditory filters centered on and above the frequencies of the middle and upper primaries, the excitation patterns for the complex tones (short dashed line) are unaffected by the addition of the noise (the solid line is on top of the short dashed line), and thus, no masking would be expected for those primaries above that caused by the neighboring component(s). Therefore, based on peripheral filtering, the noise would not be expected to reduce the peak height in the FFR spectrum at frequencies corresponding to the middle and upper primaries.

The question thus arises as to how noise can reduce the sizes of spectral components in the FFR at the frequencies of the higher primaries. In principle, when a spectral component is present in the input signal, a noise masker could reduce the size of the corresponding spectral component in the FFR due to several mechanisms. One possibility is suppression, whereby the noise reduces the BM vibration at the place responding to the primary components. A second possibility is phase-lock capture. A neuron that phase locks to a tone in the absence of noise may partially phase lock to the noise when it is added, and the phase-locked response to the tone will decrease relative to that in the absence of the noise (Marsh *et al.*, 1972). Excitation patterns are based on simultaneous masking experiments and therefore would include any effects of suppression that influence simultaneous masking. Note that for suppression to occur, a suppressor centered about two octaves below the signal frequency would need to have a level about 30 dB above the signal level to produce measurable suppression (Houtgast, 1974). Therefore, the narrowband noise masker in the present study would not be expected to suppress the middle and upper primary component in the complex. A third possibility is broadband neural inhibition, whereby inhibitory connections across frequency regions at levels above the AN lead to a reduction in phase-locked responses (Marsh *et al.*, 1972; Palmer, 1995; Wiegrebe and Meddis, 2004). Such broadband interaction above the level of the AN would not be reflected in the excitation patterns² and provides a plausible explanation for the discrepancy between masking effects

²Note that not all neurons at levels above the AN are broadly tuned (Jiang *et al.*, 1996); sharply tuned neurons also exist. Performance in tasks that require high frequency selectivity, such as detecting a tone centered in a notch noise, presumably depends on selectively monitoring the outputs of sharply tuned neurons, and so such performance would be unaffected by the presence of the broadly tuned neurons. Nevertheless, the responses of the broadly tuned neurons might contribute to the FFR.

predicted on the basis of the excitation pattern, which are more restricted in frequency, and the observed noise effects in the FFR, which are more widespread in frequency.

We next compare our results to those of some other studies, investigating frequency selectivity in the FFR. Most previous studies of the effect of noise on the FFR to tones used 500-Hz pure tone signals with noise bands centered at or above the frequency of the sinusoid (for a review, see Krishnan, 2006). For moderate signal levels, those studies reported a reduction in the FFR when the noise was within a two-octave range above the signal frequency. Huis in't Veld *et al.* (1977) also used a noise masker below the signal frequency. They presented an 8-ms 500-Hz tone at 83 dB peak equivalent SPL together with a 300-Hz wide noise band at SNRs of -3 and -13 dB. When the noise band was centered at 250 Hz (tone frequency = 1.25 times the upper edge frequency of the noise), they reported no influence on the FFR. This result contrasts with the present findings, where an effect of the noise was observed for the lower primary frequencies (primary frequencies = 1.24 and 1.43 times above the upper edge frequency of the noise for the downward- and the upward-shifted complex, respectively) and even for the upper primary frequencies (primary frequencies = 2.6 and 2.73 times above the upper edge frequency of the noise for the downward- and the upward-shifted complex, respectively). Huis in't Veld *et al.* (1977) seem to have assessed the effect of the noise bands by visual inspection of the FFR waveforms. This method would be less sensitive than computation of the magnitude spectrum of the FFR, the method used in the present study.

The noise also reduced the peak in the FFR spectrum at the QDT frequency by 4.2 dB, which is larger than expected, based on the findings of Greenberg *et al.* (1987). They measured the FFR in response to harmonics two to five of a 366-Hz F0 presented at 69 dB SPL per component. A 200-Hz wide narrowband noise, centered on 366 Hz and presented at 81 dB SPL, produced a 1-dB reduction in the FFR magnitude at F0. Greenberg *et al.* (1987) did not report whether or how the noise affected the response to the primary components in the FFR. Therefore it is possible that the QDT component in the FFR was reduced more in our study than in their study because the noise had a larger effect on the primaries themselves. This could be a consequence of the difference in stimulus parameters; the present stimuli had fewer (higher number) components in the complex tone, a higher level of the noise, and a somewhat smaller frequency ratio between the lowest harmonic in the complex and the upper edge frequency of the noise band than the stimuli used by Greenberg *et al.* (1987).

The noise bands, tone complexes, and presentation levels used in the present study were nearly identical to those used by Wile and Balaban (2007). Their objective in using the noise was to mask audible distortion products, so that the measured FFR would not reflect neural activity arising from the propagated part of the DPs. While it is likely that the noise would have masked the audible DPs (if they were audible in the first place), the present results indicate not only that the contribution of the audible/propagated part of the DPs to the observed FFR would have been very small but also that the noise would have affected the FFR response to the primaries and the QDT. Wile and Balaban (2007) did not measure the FFR in the absence of the noise.

3 Future studies—Generally, the degree to which the neural activity arising from propagated components of DPs contributes to the overall FFR, and the effects of the noise, will depend on the exact stimulus parameters. Future studies might investigate the influence of various stimulus parameters, for example, the number and the rank of harmonics and the position of the noise band, on these two effects. The results should give some further insights into the origins of the FFR.

The present results indicate that caution is needed when employing noise bands to mask audible distortion products in FFR studies, as the noise affected the response to a range of frequencies extending well beyond that expected from excitation patterns based on behavioral simultaneous masking experiments. Effectively, noise that would have negligible effects on the masking or loudness of upper primary components nevertheless reduced the FFR response to those components. Thus, unless the research question specifically investigates the effect of noise, to obtain a high SNR in the FFR to the primaries it might be preferable not to mask audible distortion products, at least when the stimulus consists mainly of resolved harmonics.

Our results are consistent with the idea (Dau, 2003) that the FFR is dominated by responses from the low-frequency tails of tuning curves of neurons with CFs well above the frequencies of the primary tones. If this applies to the QDT, then the strong component in the FFR would be markedly reduced or eliminated by a broadband noise with a level/ERB_N 10-15 dB below the level of the primary tones. However, such a noise has only a small effect on difference limens for the discrimination of F0 of complex tones with resolved components psychophysically (e.g., Gockel *et al.*, 2006). If future work shows that broadband or highpass noise does indeed markedly reduce the magnitude of the QDT in the FFR, this would support the conclusion of Gockel *et al.* (2011) that this component is not closely related to perceived pitch.

V Summary and Conclusions

The FFR was measured for two complex tones containing harmonics two, three, and four of a 300-Hz F0, with all harmonics shifted together in frequency either down or up by 56 Hz. Complex tones were presented at a level of 75 dB SPL, either in silence or in the presence of a narrowband noise, as used in previous studies to mask audible distortion products. The 85-dB SPL noise masker was 300-Hz wide and centered on 290 Hz and 310 Hz for the downward- and the upward-shifted complex tones, respectively. The effective levels of the CDT evoked by the downward-shifted complex and the QDT evoked by the downward- and the upward-shifted complex tones were estimated by comparing the magnitudes of the spectral component of the FFR at the CDT (and the QDT) frequency with that of the spectral component of the FFR for a 244-Hz sinusoid presented at various levels.

The results showed:

- (1) In agreement with previous findings, the spectral amplitude of the FFR at the CDT frequency was smaller than at the QDT frequency (i.e., the envelope-following response), the opposite to what is typically found psychophysically.

- (2) The effective CDT level in the FFR was markedly higher than that commonly estimated in psychophysical experiments, indicating strong contributions to the CDT in the FFR beyond that originating from the audible propagated component. This was also true (and even more so) for the QDT in spite of all components in the complex tones being resolved in the auditory periphery. The additional non-propagated contributions increased the effective levels of the DPs in the FFR by at least 9 and 25 dB above those estimated for the audible propagated contributions for the CDT and QDT, respectively. We argue that for these widely spaced input components, the large DPs in the FFR originate either in activity of neurons tuned to frequencies higher than the primaries or in broadly tuned neurons above the level of the AN.
- (3) The narrowband noise masker not only reduced the magnitude of the spectral component of the FFR at the CDT frequency but also had significant effects at the QDT and at all primary frequencies, for both the downward- and the upward-shifted tone complexes. We argue that the wide-ranging effect of the noise masker indicates the influence of broadly tuned neurons above the level of the AN on the FFR.

The results show that sometimes audible DPs contribute very little to the DPs observed in the FFR and that the use of a narrowband noise masker, for the purpose of masking audible DPs, can have undesired effects on the FFR over a wide frequency range.

Acknowledgments

We thank Brian Moore for helpful discussions and comments on previous versions of this paper. We would also like to thank three anonymous reviewers for helpful comments. This work was supported by Wellcome Trust Grant No. 088263.

References

- Aiken SJ, Picton TW. Envelope and spectral frequency-following responses to vowel sounds. *Hear Res.* 2008; 245:35–47. [PubMed: 18765275]
- Arnold S, Burkard R. The auditory evoked potential difference tone and cubic difference tone measured from the inferior colliculus of the chinchilla. *J Acoust Soc Am.* 1998; 104:1565–1573. [PubMed: 9745739]
- Bhagat SP, Champlin CA. Evaluation of distortion products produced by the human auditory system. *Hear Res.* 2004; 193:51–67. [PubMed: 15219320]
- Buunen TJF, Festen JM, Bilsen FA, van den Brink G. Phase effects in a three-component signal. *J Acoust Soc Am.* 1974; 55:297–303. [PubMed: 4821832]
- Carcagno S, Plack CJ. Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol.* 2011; 12:89–100. [PubMed: 20878201]
- Chertoff ME, Hecox KE, Goldstein R. Auditory distortion products measured with averaged auditory evoked potentials. *J Speech Hear Res.* 1992; 35:157–166. [PubMed: 1735965]
- Dau T. The importance of cochlear processing for the formation of auditory brain stem and frequency following responses. *J Acoust Soc Am.* 2003; 113:936–950. [PubMed: 12597187]
- Dolphin WF, Chertoff ME, Burkard R. Comparison of the envelope following response in the Mongolian gerbil using two-tone and sinusoidally amplitude-modulated tones. *J Acoust Soc Am.* 1994; 96:2225–2234. [PubMed: 7963035]
- Frisina RD. Subcortical neural coding mechanisms for auditory temporal processing. *Hear Res.* 2001; 158:1–27. [PubMed: 11506933]

- Furst M, Rabinowitz WM, Zurek PM. Ear canal acoustic distortion at $2f_1$ - f_2 from human ears: Relation to other emissions and perceived combination tones. *J Acoust Soc Am*. 1988; 84:215–221. [PubMed: 3411050]
- Galbraith GC. Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalogr Clin Neurophysiol*. 1994; 92:321–330.
- Glasberg BR, Moore BCJ, Nimmo-Smith I. Comparison of auditory filter shapes derived with three different maskers. *J Acoust Soc Am*. 1984; 75:536–544. [PubMed: 6699291]
- Glaser EM, Suter CM, Dasheiff R, Goldberg A. The human frequency-following response: Its behavior during continuous tone and tone burst stimulation. *Electroencephalogr Clin Neurophysiol*. 1976; 40:25–32. [PubMed: 55345]
- Gockel H, Moore BCJ, Plack CJ, Carlyon RP. Effect of noise on the detectability and fundamental frequency discrimination of complex tones. *J Acoust Soc Am*. 2006; 120:957–965. [PubMed: 16938983]
- Gockel HE, Carlyon RP, Mehta A, Plack CJ. The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. *J Assoc Res Otolaryngol*. 2011; 12:767–782. [PubMed: 21826534]
- Goldstein JL. Auditory nonlinearity. *J Acoust Soc Am*. 1967; 41:676–689. [PubMed: 6045077]
- Greenberg S, Marsh JT, Brown WS, Smith JC. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hear Res*. 1987; 25:91–114. [PubMed: 3558136]
- Hall JW. Auditory distortion products f_2 - f_1 and $2f_1$ - f_2 . *J Acoust Soc Am*. 1972; 51:1863–1871.
- Hartmann, WM. Signals, Sound, and Sensation. AIP; Woodbury, NY: 1997. p. 491-521.
- Houtgast, T. Lateral suppression in hearing. Ph.D. thesis, Free University of Amsterdam; 1974.
- Houtgast T. Auditory-filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker. *J Acoust Soc Am*. 1977; 62:409–415. [PubMed: 886078]
- Houtsma AJM, Goldstein JL. The central origin of the pitch of complex tones: Evidence from musical interval recognition. *J Acoust Soc Am*. 1972; 51:520–529.
- Howell, DC. Statistical Methods for Psychology. Duxbury; Belmont, CA: 1997. p. 464-466.
- Huis in't Veld F, Osterhammel P, Terkildsen K. The frequency selectivity of the 500 Hz frequency following response. *Scand Audiol*. 1977; 6:35–42. [PubMed: 594636]
- Jiang D, Palmer AR, Winter IM. Frequency extent of two-tone facilitation in onset units in the ventral cochlear nucleus. *J Neurophysiol*. 1996; 75:380–395. [PubMed: 8822565]
- Krishnan A. Human frequency-following responses to two-tone approximations of steady-state vowels. *Audiol Neurootol*. 1999; 4:95–103. [PubMed: 9892760]
- Krishnan, A. Frequency-following response. *Auditory Evoked Potentials: Basic Principles and Clinical Application*. Burkard, RF, Don, M., Eggermont, JJ., editors. Lippincott, Williams, and Wilkins; Philadelphia: 2006. p. 313-333.
- Krishnan A, Plack CJ. Neural encoding in the human brain-stem relevant to the pitch of complex tones. *Hear Res*. 2011; 275:110–119. [PubMed: 21167923]
- Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res*. 2005; 25:161–168. [PubMed: 15935624]
- Kuwada S, Anderson JS, Batra R, Fitzpatrick DC, Teissier N, D'Angelo WR. Sources of the scalp-recorded amplitude-modulation following response. *J Am Acad Audiol*. 2002; 13:188–204. [PubMed: 12025895]
- Licklider, JCR. Auditory frequency analysis. *Information Theory*. Cherry, C., editor. Academic; New York: 1956. p. 253-268.
- Marsh JT, Brown WS, Smith JC. Far-field recorded frequency-following responses: Correlates of low pitch auditory perception in humans. *Electroencephalogr Clin Neurophysiol*. 1975; 38:113–119. [PubMed: 45941]
- Marsh JT, Smith JC, Worden FG. Receptor and neural responses in auditory masking of low frequency tones. *Electroencephalogr Clin Neurophysiol*. 1972; 32:63–74. [PubMed: 4109916]
- Moore BCJ, Glasberg BR, Baer T. A model for the prediction of thresholds, loudness and partial loudness. *J Audio Eng Soc*. 1997; 45:224–240.

- Moore BCJ, Gockel HE. Resolvability of components in complex tones and implications for theories of pitch perception. *Hear Res.* 2011; 276:88–97. [PubMed: 21236327]
- Oxenham AJ, Micheyl C, Keebler MV. Can temporal fine structure represent the fundamental frequency of unresolved harmonics? *J Acoust Soc Am.* 2009; 125:2189–2199. [PubMed: 19354395]
- Palmer, AR. Neural signal processing. Hearing. Moore, BCJ., editor. Academic; Oxford: 1995. p. 75-121.
- Pandya PK, Krishnan A. Human frequency-following response correlates of the distortion product at 2F1-F2. *J Am Acad Audiol.* 2004; 15:184–197. [PubMed: 15119460]
- Patterson RD. Auditory filter shapes derived with noise stimuli. *J Acoust Soc Am.* 1976; 59:640–654. [PubMed: 1254791]
- Pressnitzer, D., Patterson, RD. Distortion products and the pitch of harmonic complex tones. *Physiological and Psychophysical Bases of Auditory Function.* Breebaart, DJ.Houtsma, AJM.Kohlrausch, A.Prijs, VF., Schoonhoven, R., editors. Shaker; Maastricht: 2001. p. 97-104.
- Rickman MD, Chertoff ME, Hecox KE. Electrophysiological evidence of nonlinear distortion products to two-tone stimuli. *J Acoust Soc Am.* 1991; 89:2818–2826. [PubMed: 1918625]
- Russo NM, Skoe E, Trommer B, Nicol T, Zecker S, Bradlow A, Kraus N. Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clin Neurophysiol.* 2008; 119:1720–1731. [PubMed: 18558508]
- Skoe E, Kraus N. Auditory brain stem response to complex sounds: A tutorial. *Ear Hear.* 2010; 31:302–324. [PubMed: 20084007]
- Smith JC, Marsh JT, Brown WS. Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalogr Clin Neurophysiol.* 1975; 39:465–472. [PubMed: 52439]
- Smooenburg GF. Audibility region of combination tones. *J Acoust Soc Am.* 1972a; 52:603–614.
- Smooenburg GF. Combination tones and their origin. *J Acoust Soc Am.* 1972b; 52:615–632.
- Wiegrefe L, Meddis R. The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus. *J Acoust Soc Am.* 2004; 115:1207–1218. [PubMed: 15058342]
- Wiegrefe L, Winter IM. Temporal representation of iterated rippled noise as a function of delay and sound level in the ventral cochlear nucleus. *J Neurophysiol.* 2001; 85:1206–1219. [PubMed: 11247990]
- Wile D, Balaban E. An auditory neural correlate suggests a mechanism underlying holistic pitch perception. *PloS ONE.* 2007; 2:e369. [PubMed: 17426817]
- Young ED, Spirou GA, Rice JJ, Voigt HF. Neural organization and responses to complex stimuli in the dorsal cochlear nucleus. *Philos Trans R Soc London Ser B.* 1992; 336:407–413. [PubMed: 1354382]
- Zwicker E. Der ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohres (The unusual amplitude-frequency characteristics of the nonlinear distortion products of the ear). *Acustica.* 1955; 5:67–74.
- Zwicker E. Different behaviour of quadratic and cubic difference tones. *Hear Res.* 1979; 1:283–292. [PubMed: 541277]
- Zwicker E. Dependence of level and phase of the $(2f_1 - f_2)$ -cancellation tone on frequency range, frequency difference, level of primaries, and subject. *J Acoust Soc Am.* 1981; 70:1277–1288.
- Zwicker E, Martner O. On the dependence of $(f_2 - f_1)$ difference tones on subject and on additional masker. *J Acoust Soc Am.* 1990; 88:1351–1358. [PubMed: 2229670]

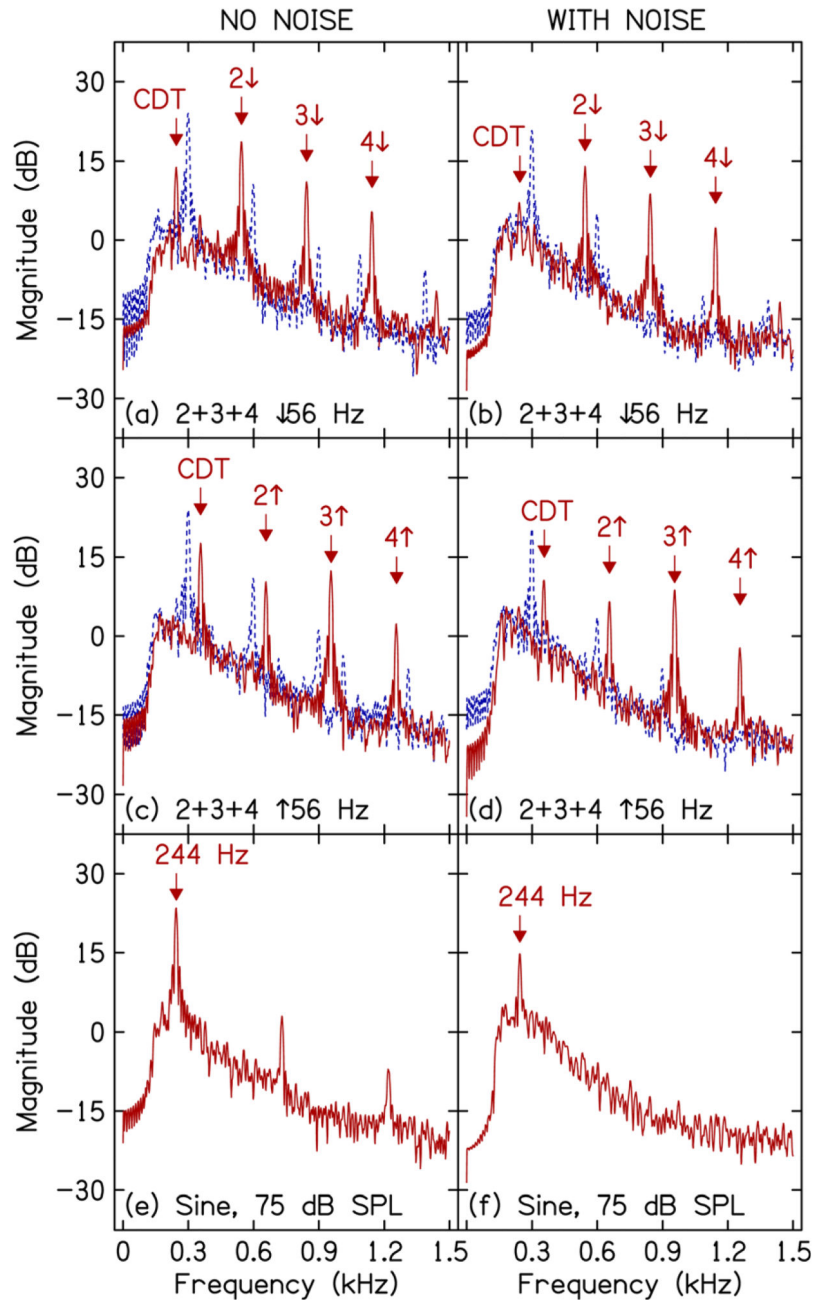


Fig. 1. (Color online) Magnitude spectra of FFRs averaged across six subjects with FFRs for the two polarities added (dashed line) or subtracted (solid line). Panels (a) and (b) show the FFR spectra for a frequency-shifted complex tone, presented at 75 dB SPL, for which harmonics 2+3+4 of a 300-Hz F₀ were shifted down by 56 Hz in the absence and presence of a narrowband noise (140–440 Hz, presented at 85 dB SPL), respectively. Panels (c) and (d) are as (a) and (b), but with the harmonics shifted up by 56 Hz and a noise in the range of 160–460 Hz. Panels (e) and (f) show the FFR spectra for a 244-Hz 75-dB SPL pure tone in the

absence and presence, respectively, of a narrowband noise (84–404 Hz) presented at 90 dB SPL.

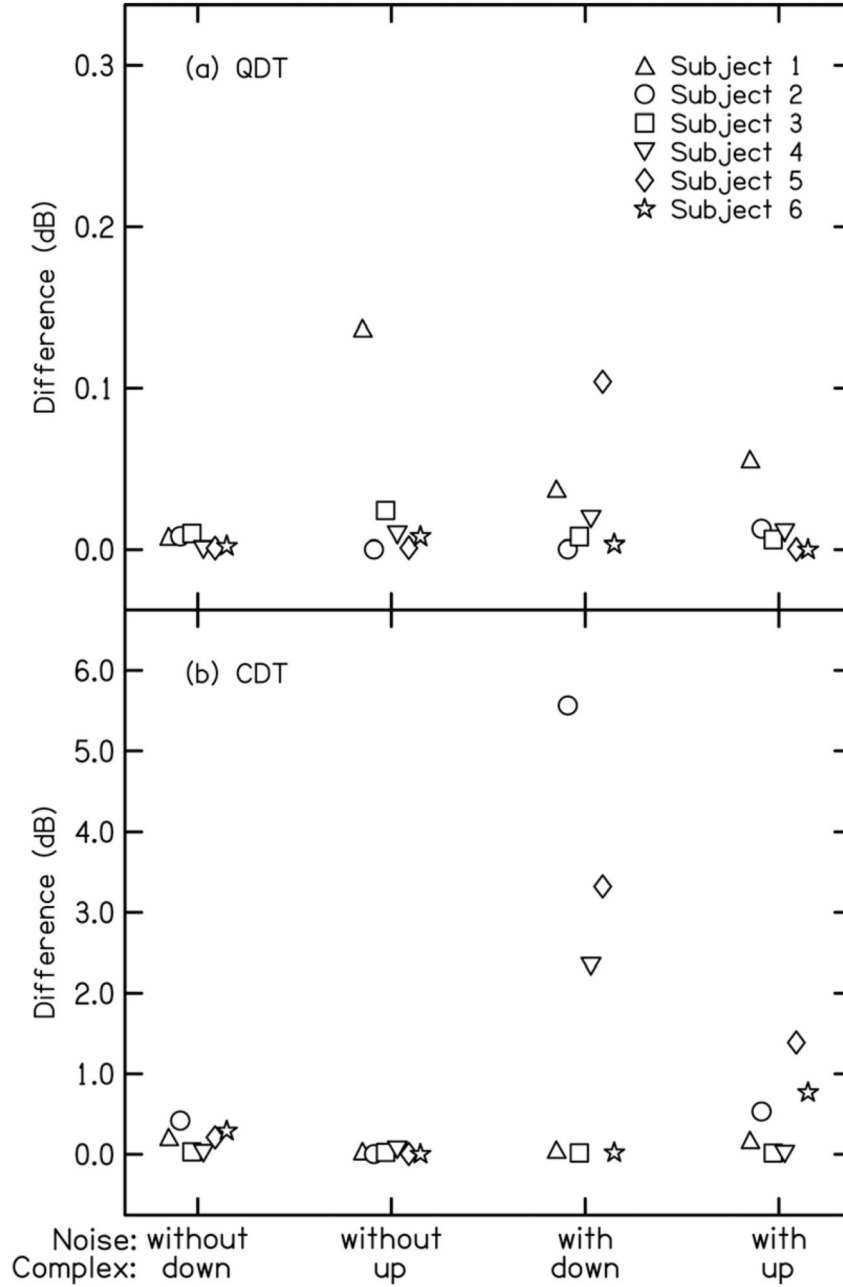


Fig. 2.

Top (a): Peak height (in dB) at the QDT frequency in the averaged spectra of the FFR waveforms for the two polarities minus the peak height at the QDT frequency in the spectrum of the addition waveform. Bottom (b): Peak height (in dB) in the FFR at the CDT frequency in the averaged spectra of the FFR waveforms for the two polarities minus the peak height at CDT in the spectrum of the subtraction waveform. The frequency-shifted complex was presented at 75 dB SPL. The 300-Hz wide noise had an rms level of 85 dB

SPL, and was centered on 290 and 310 Hz for the downward- and upward-shifted complex tones, respectively.

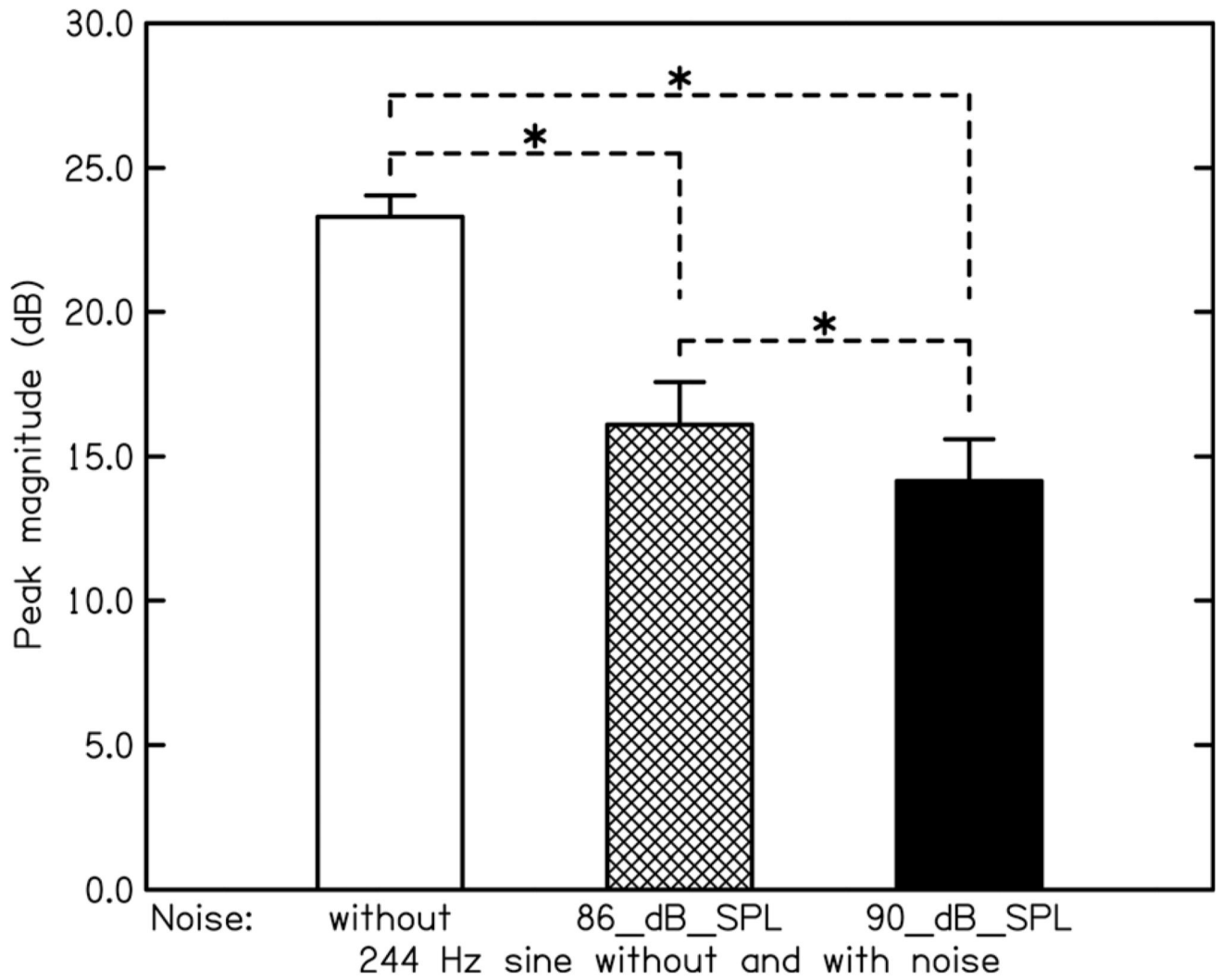


Fig. 3. Mean peak magnitude at 244 Hz (and corresponding standard error) in the spectrum of the subtraction waveform of the FFR for the 75-dB-SPL pure tone condition in the absence and in the presence of a narrowband noise (84–404 Hz) presented at levels of 86 and 90 dB SPL.

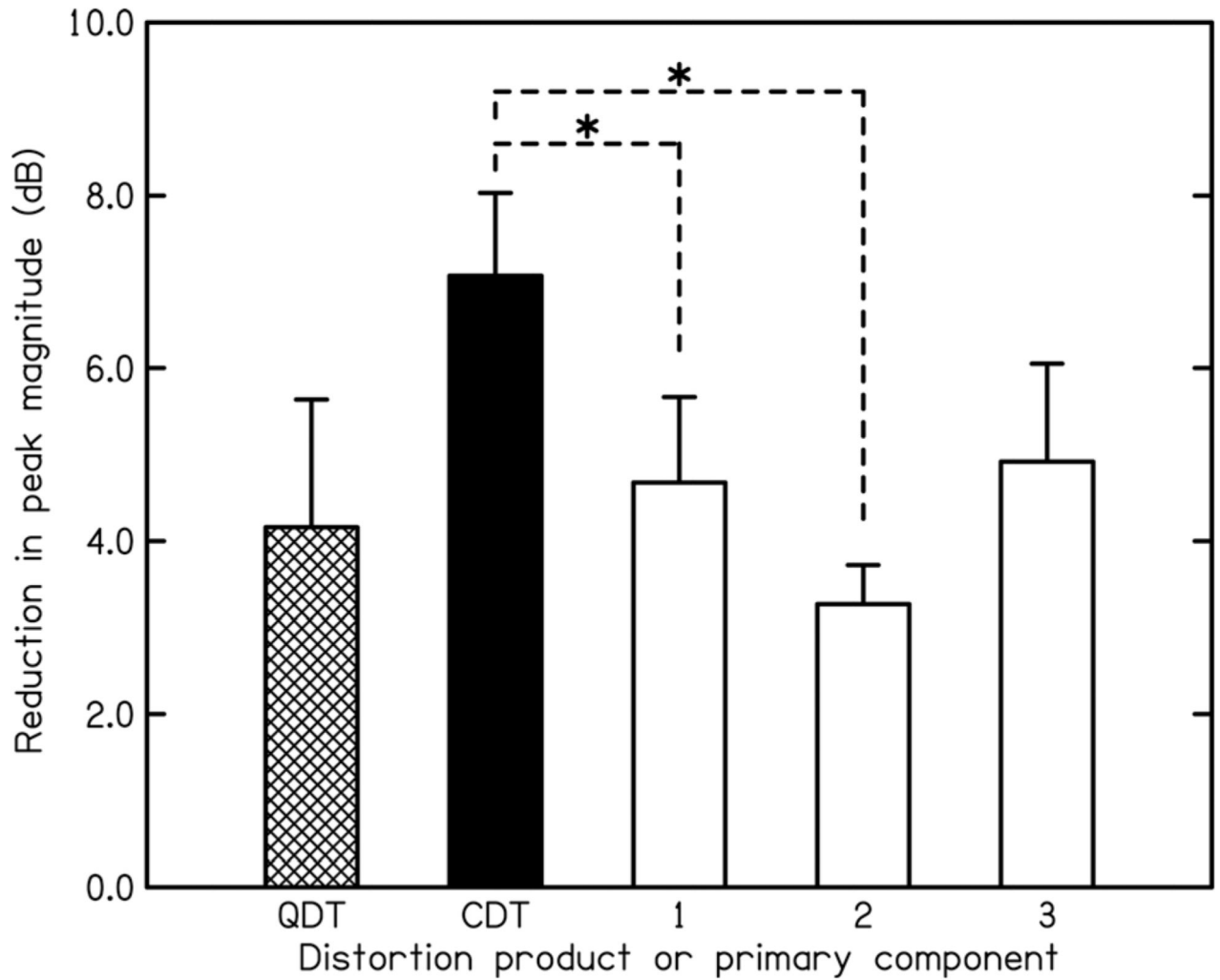


Fig. 4. Mean reduction of peak magnitudes in the FFR, due to the addition of a narrowband noise, at the QDT frequency (spectra of addition waveforms), the frequencies of the CDT, and the primary components (spectra of subtraction waveforms), averaged across conditions with the downward- and the upward-shifted complex. Complex tones were presented at 75 dB SPL. The 300-Hz wide noise had an rms level of 85 dB SPL and was centered on 290 and 310 Hz for the downward- and upward-shifted complex tones, respectively.

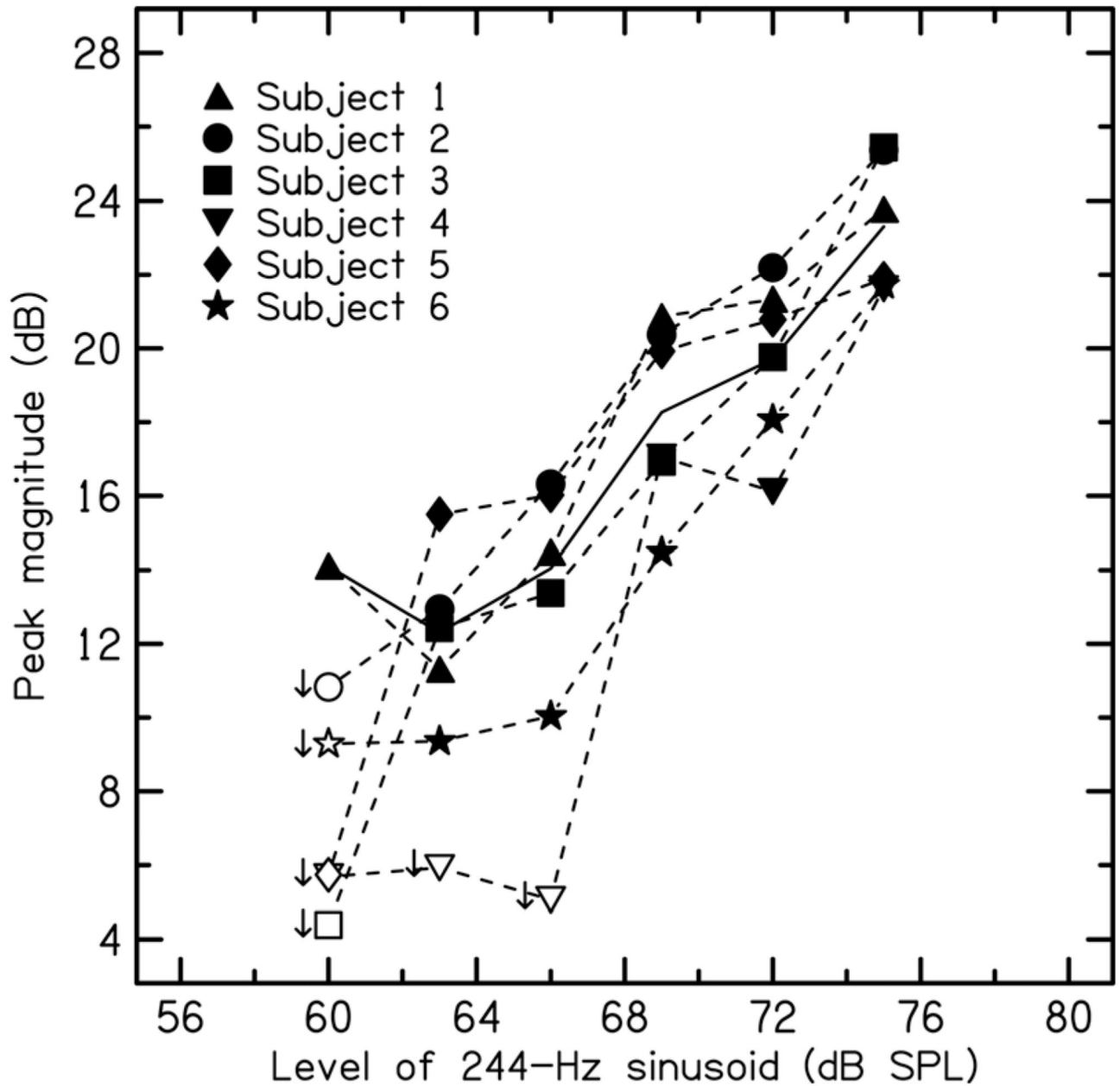


Fig. 5. Peak magnitudes at 244 Hz (spectra of subtraction FFRs) for individual subjects for the 244-Hz pure tone presented without a masker, as a function of level. Downward-pointing arrows indicate cases in which the peak value for the time interval when the tone was presented did not exceed the baseline. In these cases, the corresponding (empty) symbol gives the baseline value.

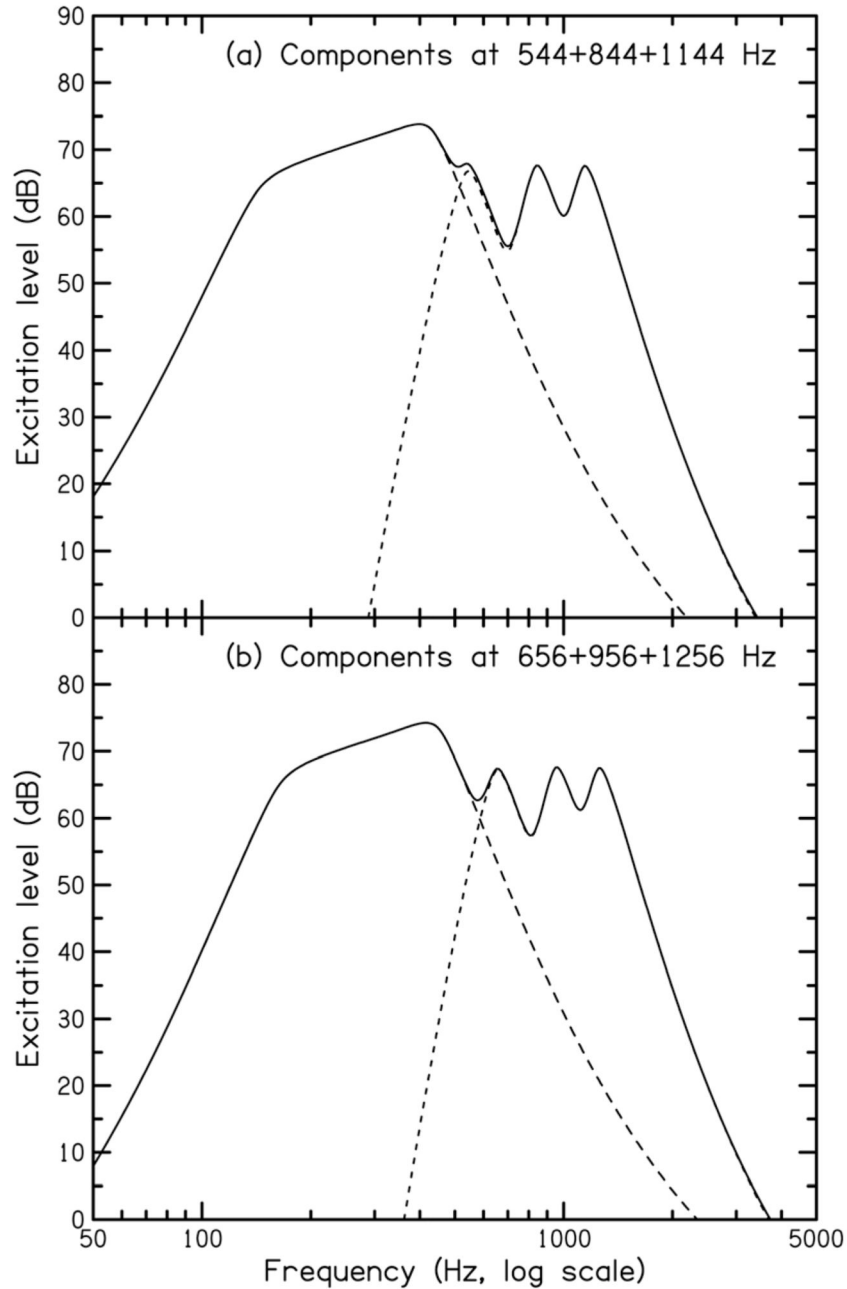


Fig. 6.

Excitation patterns for the stimuli in the complex tone conditions calculated using the model of Moore *et al.* (1997). It was assumed that the sound delivery system had a flat response at the eardrum. Excitation patterns are shown for the complex tone alone (short-dashed line), the narrowbandnoise masker alone (long-dashed line), and the two together (solid line). Excitation patterns for the downward- and upward-shifted complexes are shown at the top (a) and bottom (b), respectively.

Table I

Stimulus details for all conditions.

Condition	Components (Hz) [rms level in dB SPL]	Noise range (Hz) [rms level in dB SPL]	CDT (Hz)
Complex tone:			
(1) Downward shifted without noise	544+844+1144 (75)	–	244
(2) Downward shifted with noise	544+844+1144 (75)	140 – 440 (85)	244
(3) Upward shifted without noise	656+965+1256 (75)	–	356
(4) Upward shifted with noise	656+965+1256 (75)	160 – 460 (85)	356
Pure tone:			
(5) PT with noise	244 (75)	84 – 404 (86)	–
(6) PT with noise +	244 (75)	84 – 404 (90)	–
(7) PT without noise	244 (75)	–	–
(8) PT without noise	244 (72)	–	–
(9) PT without noise	244 (69)	–	–
(10) PT without noise	244 (66)	–	–
(11) PT without noise	244 (63)	–	–
(12) PT without noise	244 (60)	–	–

Table II

Effective levels of the cubic distortion product (CDT), defined as the input level of a 244-Hz sinusoid needed to match the spectral magnitude (at 244 Hz) in the FFR response to the downward-shifted complex. The complex tone consisted of three primaries (with frequencies of 544, 844, and 1144 Hz) with a level of 70.2 dB per component. The effective level of the CDT was estimated by linear interpolation from each subject's FFR growth function for a pure tone at the CDT frequency (see Fig. 5). The bottom line gives the mean and the standard error across subjects.

	Effective CDT level (dB SPL)	Spectral magnitude of FFR at CDT frequency for downward-shifted complex tone (dB re 0.01 μ V)
Subject 1	62.5	11.7
Subject 2	64.5	14.6
Subject 3	66.3	13.7
Subject 4	68.5	15.0
Subject 5	62.4	13.5
Subject 6	68.8	14.2
Mean	65.5 (1.2)	13.8 (0.50)

Table III

Effective level of the quadratic distortion product (QDT), defined as the input levels of a 244-Hz sinusoid needed to match the averaged spectral magnitude (in the addition waveform) of the FFR for the downward- and the upward-shifted complex tones at the QDT frequency of 300 Hz. The downward- and upward-shifted complex tones each consisted of three primaries (with frequencies of 544, 844, and 1144 Hz for the downward shift and 656, 956, and 1256 Hz for the upward shift) with a level of 70.2 dB per component. The effective level of the QDT was estimated by linear interpolation from each subject's FFR growth function for the pure tone at 244 Hz (see Fig. 5). The bottom line gives the mean and the standard error across subjects.

	Effective QDT level (dB SPL)	Averaged spectral magnitude of FFR at QDT frequency for downward- and upward-shifted complex tones (dB re 0.01 μ V)
Subject 1	68.0	18.6
Subject 2	>75.0	27.5
Subject 3	74.0	23.6
Subject 4	>75.0	22.4
Subject 5	>75.0	23.8
Subject 6	>75.0	25.4
Mean	>75.0	23.5 (1.21)