

Onset of basilar membrane non-linearity reflected in cubic distortion tone input-output functions

Robert H. Withnell, Graeme K. Yates *

The Auditory Laboratory, Department of Physiology, The University of Western Australia, Perth 6907, WA, Australia

Received 21 January 1998; revised 4 May 1998; accepted 14 May 1998

Abstract

The basilar membrane (BM) input output (I/O) function is a non-linear compressive function over much of its operating range. A low level non-compressive region with a break-point or compression threshold between 20 and 40 dB SPL has been identified. To date, no similar compression threshold in cubic distortion tone otoacoustic emission (CDT) data, which would illustrate the dependence of the CDT on BM growth, has been demonstrated. A Taylor series expansion of the outer hair cell gating function yields an amplitude term for $2f_1 - f_2$ of $p \cdot A_1^2 \cdot A_2$, where A_1 and A_2 are the displacement amplitudes of the BM for two pure tone input stimuli of levels L_1 and L_2 , p a constant. By selectively varying either L_1 or L_2 with f_2/f_1 appropriately chosen to reduce suppression effects, the CDT I/O function can be examined for deviation from the power law. In particular, if the amplitude of the CDT were dependent on BM displacement amplitude, then it should be possible by an appropriate choice of parameters to measure compression threshold. We have examined CDT I/O functions for an f_2 of 8 kHz in the guinea pig and found them to be consistent with the expected power law. With L_1 held constant, L_2 varied and $f_2/f_1 = 1.6$, a low level region with a slope of one and a compressive region with a slope of 0.14–0.27 corresponding to the analogous regions of the BM I/O function was identified, with a break-point or compression threshold of 22–33 dB SPL. © 1998 Elsevier Science B.V. All rights reserved.

Key words: Cubic distortion tone; Basilar membrane; Compression threshold; Guinea pig

1. Introduction

Distortion product otoacoustic emissions (DPOAEs) are sounds recordable in the external ear canal that are related to, but not present in, the stimulus. They are a consequence of the fact that the cochlea is a non-linear mechanical amplifier, intermodulation distortion products arising from the non-linear interaction on the basilar membrane (BM) of two simultaneously presented pure tones. The site of origin of the DPOAE is presumably near f_2 (Brown and Kemp, 1984; Kummer et al., 1995; Abdala et al., 1996; Fahey and Allen, 1997), DPOAEs being thought to be generated mostly at the place of greatest overlap of the travelling waves that result from the two pure tone stimuli.

For an input of two pure tones, the cochlea will

produce both intermodulation and harmonic distortion, in contrast to a linear system where the superposition principle holds and the output faithfully reflects the spectral input. However, distortion energy will only propagate well back to the stapes if the corresponding characteristic frequency of this energy is below its site of origin on the BM (Kirk and Yates, 1994). So whilst two pure tones together produce a range of distortion tones, those most easily measured in the ear canal include the quadratic and cubic intermodulation distortion tones. This paper will be concerned with the cubic distortion tone otoacoustic emission (CDT).

The primary source of the non-linearity in the cochlea appears to be forward transduction or mechano-electrical transduction of the outer hair cell (Patuzzi et al., 1989a,b), which, acting through the cochlear amplifier feedback loop, alters BM displacement amplitude and subsequently the displacement of the stereocilia bundle at the apex of the cell, producing a force on

* Corresponding author. Tel.: +61 (8) 9380 3321;
Fax: +61 (8) 9380 1025; E-mail: gyates@cyllene.uwa.edu.au

the gate of each transduction channel (Hudspeth, 1985). It is described by the outer hair cell gating function, a compressively non-linear function that is approximately described by a Boltzmann function – a symmetrical, saturating, sigmoidal curve (Hudspeth, 1985; Holton and Hudspeth, 1986).

There would appear to be two aspects of this forward transduction mechanism that contribute to the outer hair cell force-current function. First, there is the non-linearity associated with the statistical thermodynamics of the transduction channels: with increasing input force the output or receptor current saturates and so can no longer faithfully reflect the input. Second, a mechanical non-linearity, also due to the forward transduction process, has been identified in the deflection of the stereocilia bundle which varies the stiffness of the bundle with displacement (Howard and Hudspeth, 1988), termed gating compliance. Jaramillo et al. (1993) have considered distortion products arising from gating compliance, but either or both non-linearities could be the origin of the distortion products detected in the ear canal as otoacoustic emissions.

The cochlear amplifier appears to be governed by the forward, mechanical-to-electrical transduction stage, leading to a BM input-output (I/O) function which has three stages (see Fig. 1): a linear, low-level stage where the forward transduction is effectively linear; an intermediate, compressive stage where the forward transduction progressively saturates as the basilar membrane amplitude increases; and a high-level, linear region where the saturated forward transduction is ineffective. The stimulus intensity at which transition between the low-level and intermediate stages occurs is determined by the gain of the cochlear amplifier and the detailed properties of the forward transduction. The high-level stage has been directly observed only in preparations with a suspicion of damage (Johnstone et

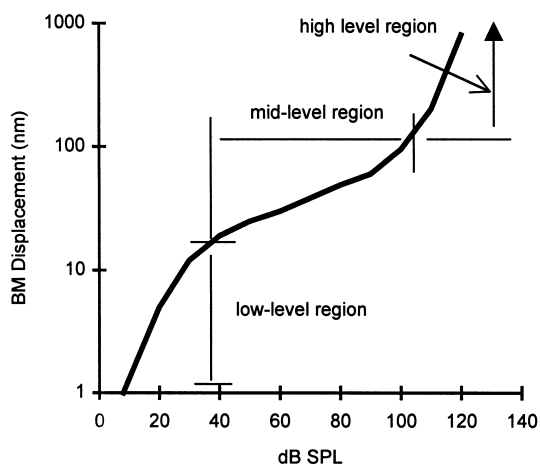


Fig. 1. Schematic of a basilar membrane input-output function measured at the characteristic frequency.

al., 1986) and has been inferred from estimates of the cochlear gain (Patuzzi et al., 1984), the transition estimated to occur at about 110 dB SPL (Ruggero et al., 1997).

The transition from the low level linear region to the compressive region of BM displacement is defined in this paper as *compression threshold* (Cooper and Yates, 1994). This point represents the onset of non-linearity in the BM I/O function. Data from BM I/O functions based on auditory nerve rate-intensity functions suggest a compression threshold of 35–40 dB SPL at a characteristic frequency of approximately 18 kHz for the guinea pig (Yates et al., 1990; Yates, 1990), while BM I/O functions measured directly using laser techniques suggest a compression threshold of approximately 20 dB SPL for characteristic frequencies of 9–10 kHz in the chinchilla (Ruggero et al., 1997) and approximately 35 dB SPL for a characteristic frequency of approximately 18 kHz in the guinea pig (Murugasu and Russell, 1996, Fig. 3).

The force function for the generation of otoacoustic emissions, the Boltzmann function, is representable as a convergent power series called a Taylor series (Marsden, 1973). With an input function of two pure tones f_1 and f_2 , the CDT ($2f_1 - f_2$) force generated at any point along the membrane is obtained with an amplitude term of $p.A_1^2.A_2$, where A_1 and A_2 are the displacement amplitudes of the BM at that point, p a constant (Sun et al., 1994a; Yates and Kirk, 1997). That is, $2f_1 - f_2$ varies as the square of A_1 multiplied by A_2 . If one substitutes L_1 and L_2 for A_1 and A_2 , where L_1 and L_2 are the stimulus intensity levels of f_1 and f_2 respectively (arguing an equivalence between stimulus level and the resultant displacement amplitude of the BM, but in fact disregarding the non-linear nature of the cochlea and the site of origin of the CDT), then the CDT for co-varying stimulus intensity levels would presumably grow at a rate on a log-log scale of three times the increase in stimulus intensity level. For L_1 or L_2 varied independently, the rate of growth of the CDT is expected to be equivalent to the increase in stimulus level when L_2 is varied and twice the increase in stimulus level when L_1 is varied.

In contrast to the predicted slope, measured CDT I/O functions obtained with co-varying stimulus intensity levels have been reported with a slope of between ~ 0.6 and ~ 1 for pure tone stimuli between 20 and 70 dB SPL in human adults (Gaskill and Brown, 1990; Popelka et al., 1993), and ~ 0.8 in the guinea pig between 35 and 65 dB SPL (Brown and Gaskill, 1990). Brown and Gaskill (1990) have considered CDT I/O functions with either L_1 or L_2 held constant, reporting slopes of 1.25 and 1.69 in humans and guinea pigs respectively for L_1 varied and 0.85 and 0.86 for L_2 varied, slopes “based on a minimum of three points ... before the response saturates” (p. 846). Suppressive effects complicate inter-

pretation of their data, this being further considered in Section 3.

A possible explanation for the discrepancy between predicted and observed slopes of CDT I/O functions is that the CDT is modified by BM displacement amplitude, such that with the onset of BM non-linearity the growth of the CDT is modified. With this hypothesis, in the low level, linear region, the CDT I/O function would therefore have a slope of 3 for co-varying stimuli, a slope of 1 for L_2 varied independently, and a slope of 2 for L_1 varied independently. In the compressive region of the BM I/O function, the growth of the CDT in all cases would be modified by BM non-linearity. The behaviour at high stimulus intensities is less clear, however, since at this intensity the forward transduction is completely saturated and so the Taylor series analysis breaks down; the CDT might not increase at all with intensity.

If the CDT follows a power law then its growth will be proportional to $A_1^2 \cdot A_2$. This study sought to examine the hypothesis that the amplitude of the CDT originating from any point on the BM is dependent only on BM displacement amplitude at that point. Ideally this should be done with a range of frequencies, but in this paper we have used only an f_2 of 8 kHz. Using a very low-noise microphone and amplifier and long averaging times, the CDT was measured as far as practicable below BM compression threshold.

CDT I/O functions were obtained for (i) L_1 and L_2 co-varied with the CDT measured over a wide range of stimulus intensities; (ii) L_1 and L_2 varied independently, measuring the CDT over a wide range of stimulus intensities; (iii) L_1 and L_2 co-varied with the addition of a suppressor tone at a frequency close to f_2 , to suppress and linearise BM vibration amplitude at the f_2 characteristic place and measure the CDT over a range of stimulus intensities.

In this paper, hereafter, we define A_1 and A_2 as the displacement amplitudes of the BM at the f_2 characteristic place (the presumed site of origin of the CDT) that result from two pure tone stimuli with stimulus levels L_1 and L_2 .

The ratio f_2/f_1 was expected to be a significant factor for these experiments due to mutual suppressive effects (Kim et al., 1980). With increasing f_2/f_1 , the suppressive effect would be expected to reduce to that of L_1 on the f_2 place, without any suppression by L_2 on f_1 . As a result, a frequency ratio of 1.6 was chosen for L_1 or L_2 varied independently to minimise suppressive effects.

2. Methods

Pigmented guinea pigs (480–800 g) were anaesthetised with Nembutal (30–35 mg/kg i.p.) and Atropine (0.06 mg i.p.), followed approximately 15 min later by Lep-

tan (0.15–0.2 ml i.m.). Neuroleptanaesthesia (Evans, 1979) was maintained using supplementary doses of Nembutal and Leptan. The guinea pigs were tracheostomised and artificially respired on Carbogen (5% CO_2 in O_2), with body (rectal) temperature maintained at 37°C. The head was positioned using a head-holder which could then be rotated for access to the ear canal. Alloferin or pancuronium (0.15 ml i.m.) was administered to reduce stapedius muscle contractions. The bulla was opened post-auricularly and a silver wire electrode placed on the round window niche for recording of the compound action potential (CAP) to monitor the condition of the cochlea. Adequate bulla ventilation was ensured by placing a plastic tube in the bulla opening. During paralysis no change in heart rate was observed.

Animals were killed with a lethal injection of Lethobarb intraperitoneally (260–325 mg Pentobarbitone Sodium) and ventilation turned off 10–30 min after the injection. Post-mortem measurements were defined as subsequent to the cessation of ventilation (applies to animals GP041 and GP044, Fig. 2).

This experiment was computer controlled with custom software and a sound-card (Crystal Semiconductor Corporation CS4231A). Ear canal sound pressures were measured with a Sennheiser MKE 2-5 Electrostatic Microphone coupled to a metal probe-tube 9 mm in length and 1.3 mm internal diameter. A 1500 Ω acoustic filter was used as an acoustic damper in the probe-tube. The probe-tube was placed approximately 2 mm into the external auditory meatus and the earphone positioned near the pinna with the position of the earphone adjusted to obtain a relatively flat ear canal sound pressure spectrum. The contralateral ear was not occluded.

The output from the probe tube microphone was amplified 20 dB, band-pass filtered (0.3–30 or 100 kHz) using two-pole Butterworth filters, and then digitised in ~ 93 ms epochs at a rate of 44.1 kHz. The epochs were averaged for a total of 10–640 s. FFTs were calculated with 256 points. Corrections for probe tube and microphone frequency responses were made.

Two digitally generated, electrically mixed, pure tone stimuli were acoustically delivered in free-field by a Foster Dynamic Earphone (Type T016H01A0000). CDT I/O functions were obtained for $f_2 = 8$ kHz and $f_2/f_1 \approx 1.2$ or 1.6. An f_2 of 8 kHz was chosen as a compromise between noise considerations for decreasing f_2 and the limit of f_2 set by the software at the time of this study. A frequency ratio of 1.2 was used where the stimulus intensity levels were co-varied. A larger f_2/f_1 ratio of 1.6 was used for either of L_1 or L_2 varied independently with the other stimulus level fixed – the larger the frequency ratio, the greater the value of L_1 that can be used without suppressing f_2 and vice-versa. By way of comparison, a frequency ratio of 1.2 for L_2 varied was also examined.

For suppression experiments, a suppressor tone with

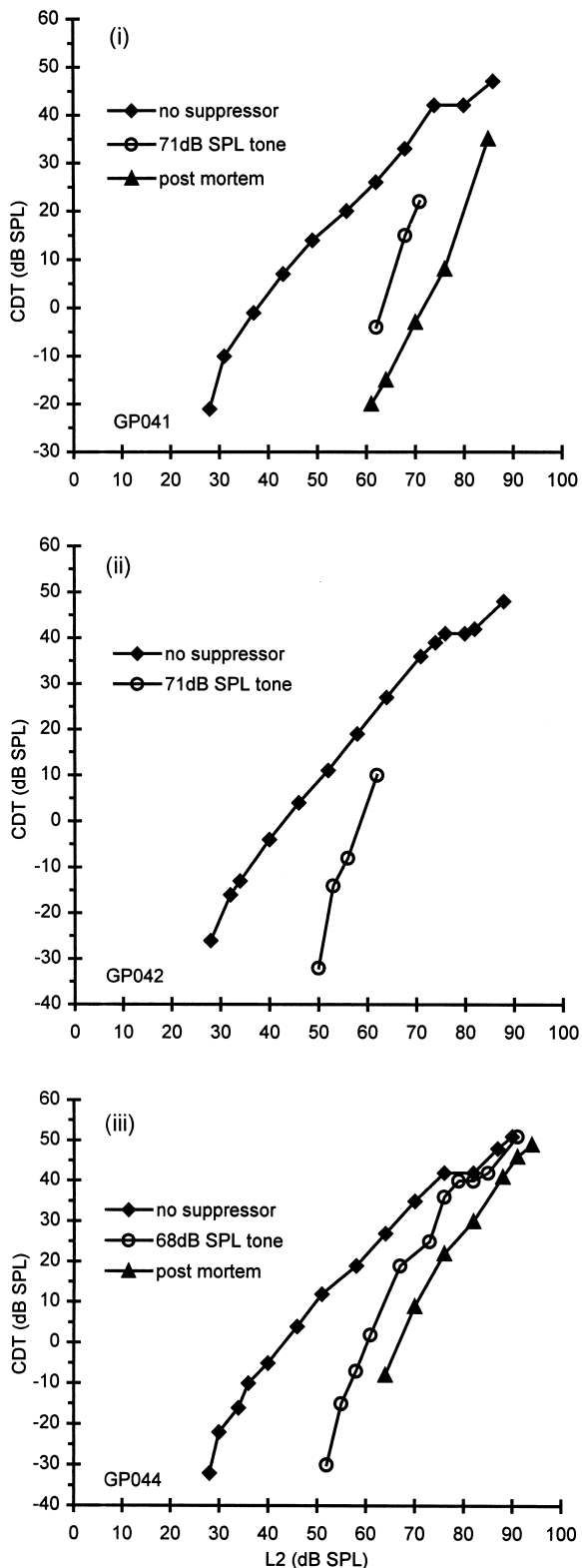


Fig. 2. CDT I/O functions for L_1 and L_2 co-varying, with an ~ 70 dB SPL 7.5 kHz suppressor tone, and approximately 90 min post mortem in (i) and (iii). $f_1 = 6.6$ kHz, $f_2 = 8$ kHz, $f_2/f_1 = 1.2$.

a frequency of 7500 Hz was generated by a Hewlett Packard 3325A Signal Generator which was then elec-

trically mixed with the output from the digital to analogue converter and delivered via the Foster earphone.

The care and use of animals reported on in this study was approved by the Animal Experimentation Ethics Committee of the University of Western Australia (approval number UWA 91/97), and all procedures conformed with the Code of Practice of the National Health and Medical Research Council of Australia.

3. Results

3.1. Co-varying L_1, L_2

Considering first CDT I/O functions obtained for L_1 and L_2 co-varying, Fig. 2 shows results obtained from three animals. Fig. 2 (i, iii) shows I/O functions obtained with and without a suppressor tone, and post mortem; in Fig. 2 (ii) there is no post-mortem result. The slopes of the CDT I/O functions (with no suppressor tone) vary from 1.2 to 1.4 over the range $30 \leq L_2 \leq 75$ dB SPL. This region could correspond to the mid level compressive region of the BM I/O function. The CDT I/O function for L_1 and L_2 co-varying however is complicated by mutual suppressive effects occurring with increasing L_1 , L_2 (Kim et al., 1980) and the CDT presumably being a sum of vector components (Zwicker, 1980).

For $L_2 < 30$ dB SPL for each of the three animals there is evidence, albeit based on one data point, of the slope being much greater than the 1.2–1.4 observed for $L_2 > 30$ dB SPL. This increase in slope is consistent across animals and would be consistent with the sound pressure level at which BM growth becomes compressive. Note that for Fig. 2 (ii, iii) the data point for the lowest stimulus level actually represents the noise floor with no CDT identified, however this clearly indicates a significant change in the slope as the actual CDT value would be of the order of, or less than, this noise floor value.

The CDT I/O functions in Fig. 2 essentially show CDT amplitude to continue to increase up to a stimulus level of ~ 90 dB SPL. This is consistent with a high level transition point for BM I/O functions in the chinchilla of the order of 110 dB SPL (Ruggero et al., 1997). A non-monotonicity is evident at ~ 75 dB SPL, the origin of non-monotonicities in the CDT I/O function having been suggested as being due to phase interactions between the CDT vector components (Zwicker, 1980). He and Schmiedt (1993) have suggested that non-monotonicities and overall variability in the CDT I/O function in humans is associated with the behaviour of CDT fine structure with level, however the CDT in the guinea pig does not demonstrate the same fine structure (Brown and Gaskill, 1990).

With the addition of an ~ 70 dB SPL suppressor

tone with a frequency just below f_2 , the CDT I/O function shifts to the right with an increase in slope to a value of ~ 3 for $L_2 < 70$ dB SPL. A slope of 3 is as predicted by simple non-linear analysis, with the active process inhibited. This is consistent with Sun et al. (1994b)'s observation that the addition of a level saturating non-linearity can modify the slope of the CDT from three to approximately one, since suppression effectively removes the compression associated with the active process (Ruggero et al., 1992). Inhibition of the active process by suppression is analogous to Mills and Rubel (1994) describing reducing the endocochlear po-

tential and obtaining a CDT I/O function which they termed as having a passive source.

Post-mortem data in Fig. 2 (i, iii) show the I/O function to shift further to the right with an increase in slope, illustrating the physiological nature of the data.

Instrumental distortion will have a slope of three for stimulus levels that are co-varied. I/O functions with the addition of a suppressor tone are not contaminated by instrumental distortion though as the post-mortem responses provide an upper limit to such distortion. Independent measures of instrumental distortion in a hard-walled cavity of dimensions commensurate with

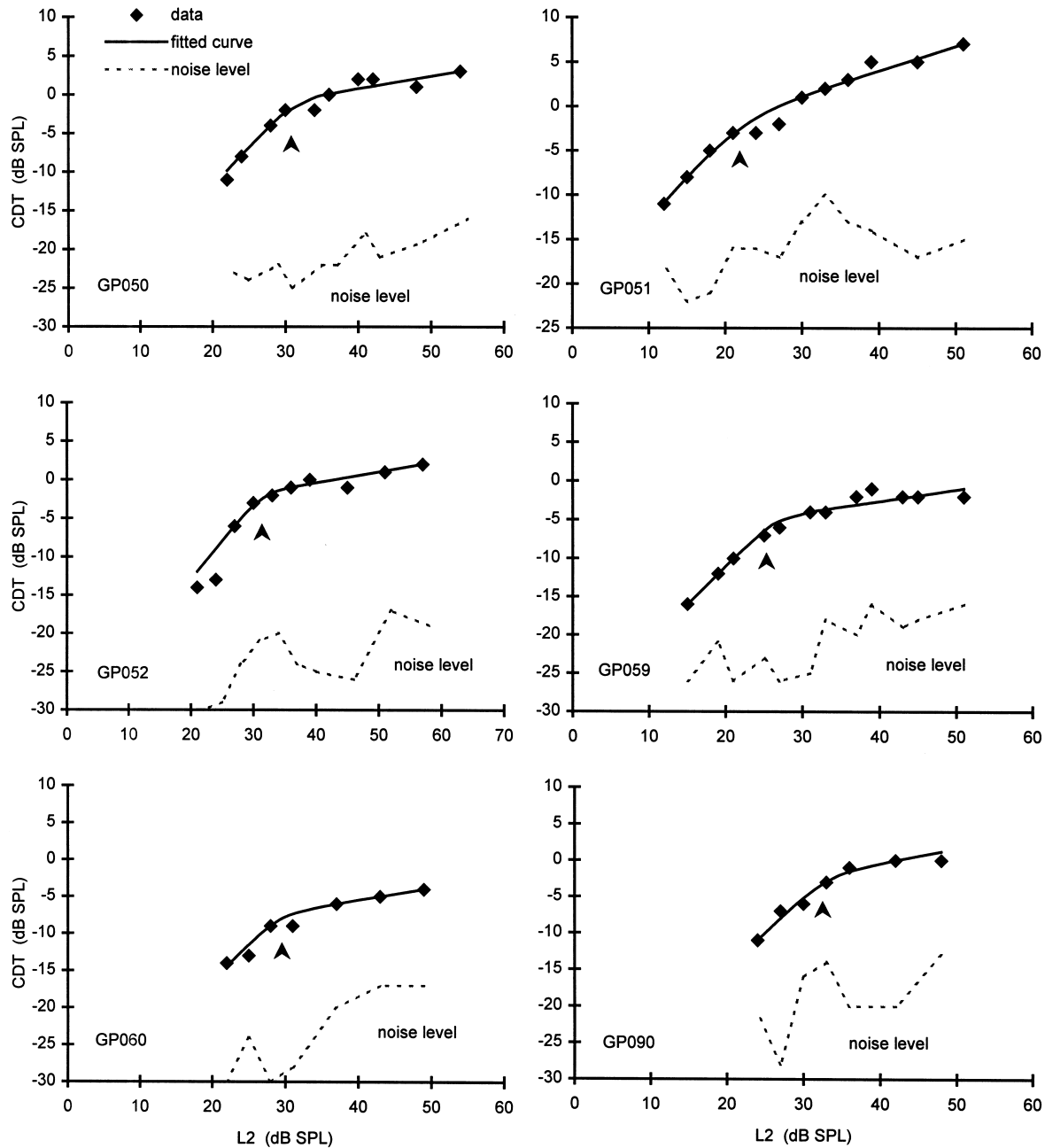


Fig. 3. CDT I/O functions for L_1 fixed ($51 \text{ dB SPL} \leq L_1 \leq 56 \text{ dB SPL}$), L_2 varied. $f_1 = 5 \text{ kHz}$, $f_2 = 8 \text{ kHz}$, $f_2/f_1 = 1.6$. A curve fitted to each data series is given by a solid line. Arrowheads indicate compression threshold. The noise level in each case is indicated by a dashed line.

the guinea pig ear canal suggest post-mortem responses not to be contaminated by such distortion (for stimulus levels of 85 dB SPL cubic instrumental distortion was less than -4 dB SPL).

3.2. L_2 varied independently

Fig. 3 provides six examples of CDT I/O functions with L_1 fixed and L_2 varied, L_1 having a value between 51 and 56 dB SPL. Fitted to the data in each example is a curve described by the equation

$$\text{CDT} = c \times \left[\frac{A_3^{(1/A_4-1)} \times p^{(1/A_4)}}{A_3^{(1/A_4-1)} + p^{(1/A_4-1)}} \right]^{A_4} \quad (1)$$

where c is a scaling factor, A_3 is the sound pressure at which the BM becomes non-linear, A_4 is the slope of the BM I/O curve in its compressive region, p is the stimulus sound pressure (in Pascals).

This equation represents the BM I/O curve derived from auditory nerve rate-intensity functions (Yates, 1991) and so represents BM displacement measured at a single point on the BM. Because CDT data presumably represent a vector sum of components arising from many different points along the BM (Zwicker, 1980), the qualifier to the use of Eq. 1 to fit CDT data is that it will only be an approximation. However, it can be seen that up to stimulus levels of L_2 of approximately 50 dB SPL, the curve fit in each example supports the hypothesis of the CDT being proportional to BM displacement amplitude. Above 50–60 dB SPL, suppressive effects complicate the interpretation (Arthur et al., 1971; Schmiedt, 1982; Delgutte, 1990), so points above ~ 50 dB SPL were not included for curve fitting.

Also shown in Fig. 3 are noise levels to give some indication of the signal to noise ratios of the CDTs measured. The noise level in each case was determined by calculating the average pressure of four points in the FFT (two points either side of the $2f_1-f_2$ point) and then calculating the sound pressure level of this average pressure value.

Table 1 shows CAP threshold, compression threshold and A_4 for each of the six animals in Fig. 3. Compression threshold values range from 22 to 33 dB SPL. The difference between CAP threshold and compression threshold is ≤ 6 dB in all cases, consistent with the

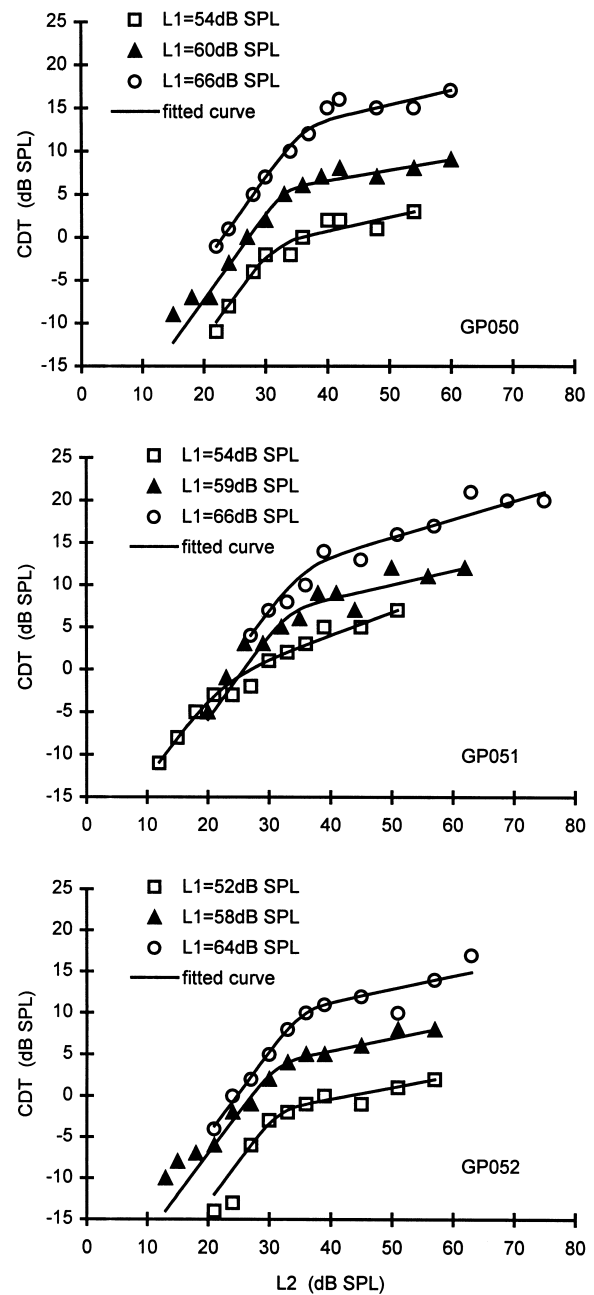


Fig. 4. CDT I/O functions with L_1 fixed, L_2 varied, for three different levels of L_1 . $f_1 = 5$ kHz, $f_2 = 8$ kHz, $f_2/f_1 = 1.6$. A curve fitted to each data series is given by a solid line.

Table 1

CAP threshold, compression threshold with L_1 fixed, L_2 varied, and A_4 for each of the six animals in Fig. 3

Animal	CAP threshold (dB SPL)	Compression threshold (dB SPL)	A_4
GP050	33	31	0.16
GP051	26	22	0.27
GP052	29	31	0.14
GP059	31	26	0.15
GP060	30	29	0.16
GP090	39	33	0.20

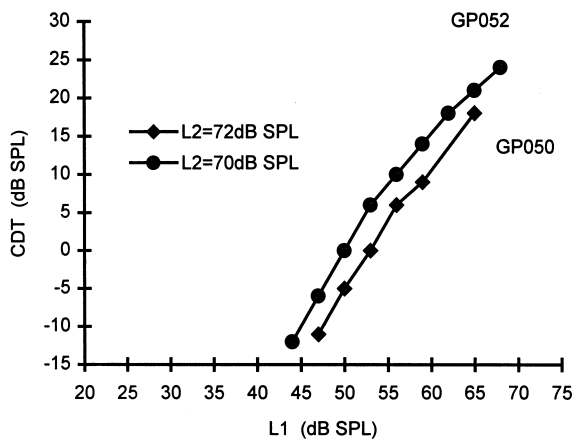


Fig. 5. CDT I/O functions with L_2 fixed ($L_2 \approx 70$ dB SPL), L_1 varied. $f_1 = 5$ kHz, $f_2 = 8$ kHz, $f_2/f_1 = 1.6$.

observation from neural data that CAP threshold occurs close to compression threshold level (Yates et al., 1990). A_4 , which represents the slope of the BM I/O curve in the compressive region, ranges from 0.14 to 0.27, in good agreement with directly measured values of 0.2–0.3 (Robles et al., 1986; Yates et al., 1990; Ruggero et al., 1997).

Fig. 4 gives three examples of CDT I/O functions obtained for L_2 varied independently, with successively increasing values of L_1 . For L_1 above ~ 60 – 65 dB SPL, it would appear that L_1 is beginning to suppress f_2 as the I/O functions move to the right, i.e. compression threshold is observed to increase. Compression threshold values from Fig. 4 are given in Table 2.

3.3. L_1 varied independently

Fig. 5 shows CDT I/O functions from two animals for L_1 varied, $L_2 \approx 70$ dB SPL. For $L_1 < 55$ dB SPL the slope of both functions is 2. For $L_1 > 56$ dB SPL both functions have a slope of ~ 1.2 . A slope of 2 below the respective ‘observed’ transition points (56 and 53 dB SPL respectively) is consistent with the expected power law growth. Now for L_1 varied, it is the growth of A_1 that determines the I/O function: at the f_2 site, compression threshold, if present, will occur at a much higher value of L_1 than that observed for L_2 in Fig. 4. In Fig. 5, it is probable that compression threshold has not been reached but in fact that the change in slope from 2 to 1.2 is due to suppression of f_2 by L_1 . This interpretation is supported by Fig. 7 of Ruggero et

al. (1997) where velocity-intensity functions of BM responses to tones with frequencies equal to and lower than a characteristic frequency of 10 kHz were presented for the chinchilla: compression thresholds of 20 dB SPL at 10 kHz, 30 dB SPL at 9 kHz, 50 dB SPL at 8 kHz, and no compression threshold up to 90 dB SPL at 7 kHz are suggested. Thus for a frequency ratio of 1.6, corresponding to an f_1 of 5 kHz, no compression threshold should be evident in the CDT I/O function up to $L_1 = 90$ dB SPL.

3.4. Interpretation for L_1 , L_2 co-varying

Armed with estimates of compression thresholds from BM I/O data from Ruggero et al. (1997), consider the CDT I/O functions for co-varying stimuli (Fig. 2). With $f_2/f_1 = 1.2$, assuming an f_2 site of origin, the 8 kHz tone (f_2) has a compression threshold of ~ 20 dB SPL whilst the 6.6 kHz tone (f_1) has a compression threshold of ~ 45 dB SPL at the f_2 place, and furthermore the growth of the 6.6 kHz tone in the compressive region at the f_2 or 8 kHz site will have a slope greater than 0.2, i.e., 0.5–0.6. So for co-varying stimuli less than 20 dB SPL, the slope should be 3, from 20 to 45 dB SPL the slope should ~ 2.2 , and above 45 dB SPL the slope should be about $(0.5 \times 2 + 0.2) = 1.2$. These estimates do not consider suppression effects, nor the hypothesis that the CDT is presumably a sum of vector components and so with spread of the CDT excitation pattern with increasing stimulus level, phase interactions become more significant. Thus predicted slopes will not necessarily match those observed in Fig. 2, and so it is difficult to interpret CDT I/O functions for co-varying stimuli.

3.5. L_2 varied with $f_2/f_1 = 1.2$ versus 1.6

Fig. 6 shows CDT I/O functions obtained with $f_2/f_1 = 1.2$ with L_2 varied ($L_1 = 40, 47, 55$ dB SPL) and for $f_2/f_1 = 1.6$ with $L_1 = 57$ dB SPL, L_2 varied. Compression threshold is 24 dB SPL in Fig. 6 (ii). Applying the curve fitting equation to Fig. 6 (i) for $L_1 = 40$ dB SPL and only including data points with $L_2 \leq 37$ dB SPL (to avoid suppressive effects), a compression threshold of 26 dB SPL is obtained, in good agreement with the 24 dB SPL obtained in Fig. 6 (ii). With $L_1 > 40$ dB SPL in Fig. 6 (i), compression threshold shifts upwards in value, L_1 presumably suppressing at the f_2 site above 40

Table 2

Compression thresholds (dB SPL) for increasing L_1 , with L_1 fixed, L_2 varied, for the three animals in Fig. 4

Average L_1 (dB SPL)	Animal GP050	Animal GP051	Animal GP052
53	31	22	31
59	33	33	31
65	36	35	35

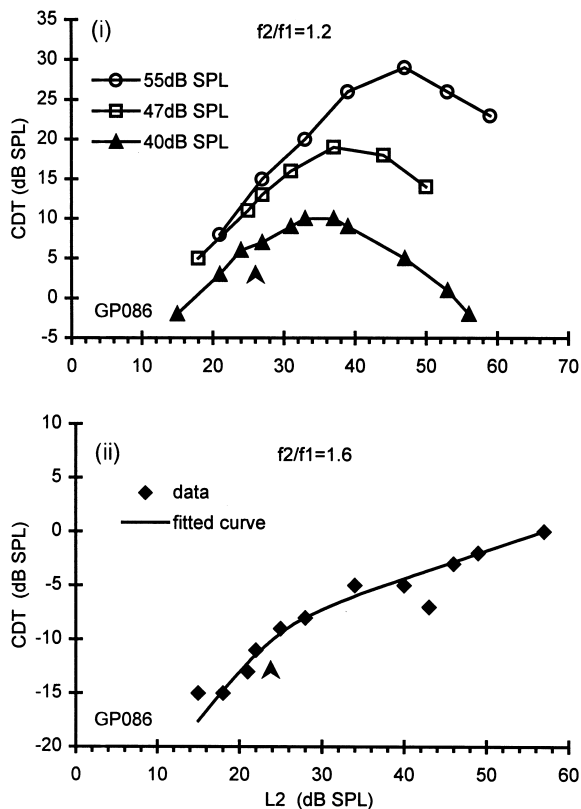


Fig. 6. CDT I/O functions (i) with $f_2/f_1 = 1.2$ with L_1 fixed and L_2 varied ($L_1 = 40, 47, 55$ dB SPL), $f_2 = 8$ kHz, and (ii) with $f_2/f_1 = 1.6$ with $L_1 = 57$ dB SPL, L_2 varied, $f_2 = 8$ kHz. Arrowheads indicate compression threshold.

dB SPL. As L_2 is further increased in value the I/O functions roll over, evidence for suppression of the f_1 primary. The I/O functions for both $L_1 = 40$ dB SPL and $L_1 = 47$ dB SPL indicate a suppression threshold of 37 dB SPL. For $L_1 = 55$ dB SPL, the f_2 place also appears to be suppressed as the roll-over point is raised to 47 dB SPL. These values are in good agreement with suppression threshold estimates from cat and gerbil neural data (Arthur et al., 1971; Schmiedt, 1982; Delgutte, 1990).

It is evident that an estimate of compression threshold may be obtained from CDT data with $f_2/f_1 = 1.2$ (provided compression threshold is sufficiently below the L_2 suppression threshold level of f_1), but suppressive effects hide the dependence of the growth of the CDT on BM displacement amplitude in the region of compressive BM growth. This dependence is better illustrated with a larger frequency ratio of 1.6.

4. Discussion

Cubic distortion tone I/O functions reported in this paper suggest that the CDT does follow the expected power law with its growth proportional to $A_1^2 \cdot A_2$,

where A_1 and A_2 are the displacement amplitudes of the BM at the f_2 characteristic place, although in general A_1 and A_2 are associated with the CDT force generated at many points on the BM. For a site of origin for the CDT near f_2 , A_1 and A_2 , or more exactly $A_1^2 \cdot A_2$, presumably only becomes significant near f_2 .

For L_2 varied independently (Fig. 3), a low level linear region and a mid-level compressive region is suggested which would correspond to the BM I/O function measured at the characteristic frequency; compression thresholds of 22–33 dB SPL were obtained, in reasonable agreement with previous estimates based on BM I/O functions (Yates et al., 1990; Yates, 1990; Murugasu and Russell, 1996; Ruggero et al., 1997). A compression threshold of 22–33 dB SPL suggests a CDT site of origin close to the characteristic place for f_2 , as has been previously suggested by Anderson (1980), and is consistent with model data from Yates and Kirk (1997). If the CDT site of origin were basal to the peak of the f_2 travelling wave, then the compression threshold for CDTs would be greater than for the BM due to the greater stimulus level required to saturate the BM at that point.

Good agreement was observed for CDT I/O functions obtained with L_2 varied with the fitted curves using Eq. 1. Now as this equation represents BM displacement measured at the characteristic frequency on the BM, it is evident that for low level stimuli, the CDT behaves as if it originates at a single characteristic place. Since the region of interaction on the BM between two travelling waves that results from two pure tone stimuli can only be near the f_2 place, and since f_2 is the component being altered, the f_2 place must be this characteristic place.

It has been suggested that there are two ‘sources’ of CDT – an ‘active’ source dominant at low stimulus intensity levels and a ‘passive’ source dominant at high stimulus intensity levels (Whitehead et al., 1990; Whitehead et al., 1992a,b; Mills and Rubel, 1994). An alternative explanation, supported in this paper, is that there is only one ‘source’, the non-linearity inherent in mechano-electrical transduction, and the amplitude of the CDT originating from any point on the BM is dependent only on BM displacement amplitude at that point. A viewpoint of only one source is supported by Mills (1997).

It is thought that CDTs are produced predominantly by the maximum overlap of two travelling waves on the BM. This maximum overlap of the travelling waves occurs essentially at the f_2 place for low level stimuli. In presenting data with L_2 varied to identify compression threshold, it has been pre-supposed that L_1 can be chosen to interact with L_2 without first suppressing at the f_2 place. It may be that the measurement of CDT is evidence for L_1 suppressing at the f_2 place. The agreement between CDT I/O functions with L_2 varied and

fitted curves described by Eq. 1 (Fig. 3) would seem to argue for the former.

In this paper, we have not considered the suppressive effects of the stimulus tones at the $2f_1-f_2$ site; Giguère et al. (1997) recently considered the suppressive effect of L_1 on the perceived cubic distortion tone, demonstrating that the effect reduces with increasing f_2/f_1 . For L_2 varied independently with an f_2/f_1 of 1.6, $f_2=8$ kHz, $f_1=5$ Hz and $2f_1-f_2=2$ kHz, suppression of $2f_1-f_2$ by L_1 with $L_1<60$ dB SPL should be non-significant. We have considered the suppressive effects of L_1 on f_2 and L_2 on f_1 , arguing for a greater f_2/f_1 to reduce suppression and so reveal the dependence of CDT amplitude on BM displacement amplitude. For the CDT in humans, Brown et al. (1996) and Brown and Gaskell (1996) suggest that both the f_2 and $2f_1-f_2$ sites contribute to the resultant CDT in some cases, with the f_2 element being the dominant component. This is the presumed source of the CDT fine structure seen in the human, however the guinea pig does not demonstrate the same fine structure (Brown and Gaskell, 1990) and so would not appear to have a secondary emission from the $2f_1-f_2$ site.

It is likely that the BM I/O function compression threshold for humans occurs at approximately 20 dB SPL at 1 kHz, based on psychoacoustic findings (Vie-meister and Bacon, 1988). A compression threshold of ≤ 20 dB SPL would explain why there has been no evidence for a cubic power law from CDT I/O functions in humans (He and Schmiedt, 1993). Cubic distortion tone I/O functions have invariably been obtained with the stimulus levels co-varying, and never extending significantly in stimulus level below 20 dB SPL. Human CDTs at such levels would be difficult to identify due to the poor signal to noise ratio.

Cubic distortion tones in humans appear to be smaller in amplitude than that found in non-human mammals. For instance, for equi-level stimuli and $f_2/f_1=1.2$, the amplitude of the CDT is some 60 dB below the stimulus level in humans but only 40 dB in guinea pigs. One possible explanation for this is the greater frequency selectivity in humans (Fay, 1988) and hence reduced overlap of the excitation patterns that result from the two pure tone stimuli, although this might be balanced to some extent by reduced suppressive effects. Suppressive effects will also be determined by the frequency selectivity of the BM.

Cubic distortion tone otoacoustic emissions have found widespread application in the evaluation of human auditory function. The focus in their application however has been on the amplitude of the CDT for fixed level stimuli with a pass/fail criteria (e.g. Stover et al., 1996; Gorga et al., 1996). Compression threshold provides a value that is closely related to CAP threshold in cochleas with normal auditory sensitivity. It awaits further studies though to determine whether

auditory sensitivity in pathologically altered cochleas can be predicted on the basis of compression threshold estimates from CDTs, or indeed if such a technique has any clinical utility.

Acknowledgments

We wish to thank Foster Electric Company of Japan for generously supplying the earphones used in this study. We also wish to thank Mr Greg Nancarrow for invaluable technical advice and support and Dr Des Kirk for discussions in respect of this paper. This work was supported by the Australian National Health and Medical Research Council, Project Grant 960566, and by the University of Western Australia. R.H.W. is the recipient of a postgraduate scholarship from the Department of Physiology. Parts of this paper were presented at the Association for Research in Otolaryngology Mid-winter Meeting, February 1998.

References

- Abdala, C., Sininger, Y.S., Ekelid, M., Zeng, F., 1996. Distortion product otoacoustic emission suppression tuning curves in human adults and neonates. *Hear. Res.* 98, 38–53.
- Anderson, S.D., 1980. Some ECMR properties in relation to other signals from the auditory periphery. *Hear. Res.* 2, 273–296.
- Arthur, R.M., Pfeiffer, R.R., Suga, N., 1971. Properties of two-tone inhibition in primary auditory neurons. *J. Physiol.* 212, 593–609.
- Brown, A.M., Gaskell, S.A., 1990. Measurement of acoustic distortion reveals underlying similarities between human and rodent mechanical responses. *J. Acoust. Soc. Am.* 88, 840–849.
- Brown, A.M., Gaskell, S.A., 1996. Suppression of human acoustic distortion product: dual origin of $2f_1-f_2$. *J. Acoust. Soc. Am.* 100, 3268–3274.
- Brown, A.M., Kemp, D.T., 1984. Suppressibility of the $2f_1-f_2$ stimulated acoustic emissions in gerbil and man. *Hear. Res.* 13, 29–37.
- Brown, A.M., Harris, F.P., Beveridge, H.A., 1996. Two sources of acoustic distortion products from the human cochlea. *J. Acoust. Soc. Am.* 100, 3260–3267.
- Cooper, N.P., Yates, G.K., 1994. Nonlinear input-output functions derived from the responses of guinea-pig cochlear nerve fibres: Variations with characteristic frequency. *Hear. Res.* 78, 221–234.
- Delgutte, B., 1990. Two-tone rate suppression in auditory nerve fibres: Dependence on suppressor frequency and level. *Hear. Res.* 49, 225–246.
- Evans, E.F., 1979. Neuroleptanaesthesia: an ideal anaesthetic procedure for long-term physiological studies of the guinea pig cochlea. *Acta Otolaryngol.* 105, 185–186.
- Fahey, P.F., Allen, J.B., 1997. Measurement of distortion product phase in the ear canal of the cat. *J. Acoust. Soc. Am.* 102, 2880–2891.
- Fay, R.R., 1988. Comparative psychoacoustics. *Hear. Res.* 34, 295–306.
- Gaskell, S.A., Brown, A.M., 1990. The behaviour of the acoustic distortion product, $2f_1-f_2$, from the human ear and its relation to auditory sensitivity. *J. Acoust. Soc. Am.* 88, 821–839.
- Giguère, C., Smoorenburg, G.F., Kunov, H., 1997. The generation of psychoacoustic combination tones in relation to two-tone suppression effects in a computational model. *J. Acoust. Soc. Am.* 102, 2821–2830.

- Gorga, M.P., Stover, L., Neely, S.T., Montoya, D., 1996. The use of cumulative distributions to determine critical values and levels of confidence for clinical distortion product otoacoustic emission measurements. *J. Acoust. Soc. Am.* 100, 968–977.
- He, N., Schmiedt, R.A., 1993. Fine structure of the $2f_1 - f_2$ acoustic distortion product: Changes with primary level. *J. Acoust. Soc. Am.* 94, 2659–2669.
- Holton, T., Hudspeth, A.J., 1986. The transduction channel of hair cells from the bull-frog characterised by noise analysis. *J. Physiol.* 375, 195–228.
- Howard, J., Hudspeth, A.J., 1988. Compliance of the hair bundle associated with gating of mechano-electrical transduction channels in the bullfrog's saccular hair cell. *Neuron* 1, 189–199.
- Hudspeth, A.J., 1985. The cellular basis of hearing: The biophysics of hair cells. *Science* 230, 745–752.
- Jaramillo, F., Markin, V.S., Hudspeth, A.J., 1993. Auditory illusions and the single hair cell. *Nature* 364, 527–529.
- Johnstone, B.M., Patuzzi, R., Yates, G.K., 1986. Basilar membrane measurements and the travelling wave. *Hear. Res.* 22, 147–153.
- Kim, D.O., Molnar, C.E., Matthews, J.W., 1980. Cochlear mechanics: Nonlinear behaviour in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear canal sound pressure. *J. Acoust. Soc. Am.* 67, 1704–1721.
- Kirk, D.L., Yates, G.K., 1994. Evidence for electrically evoked travelling waves in the guinea pig cochlea. *Hear. Res.* 74, 38–50.
- Kummer, P., Janssen, T., Arnold, W., 1995. Suppression tuning characteristics of the $2f_1 - f_2$ distortion-product otoacoustic emission in humans. *J. Acoust. Soc. Am.* 98, 197–210.
- Marsden, J.E., 1973. *Basic Complex Analysis*. W.H. Freeman and Co., San Francisco, CA.
- Mills, D.M., 1997. Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones. *J. Acoust. Soc. Am.* 102, 413–429.
- Mills, D.M., Rubel, E.W., 1994. Variation of distortion product otoacoustic emissions with furosemide injection. *Hear. Res.* 77, 183–199.
- Murugasu, F., Russell, I.J., 1996. The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea. *J. Neurosci.* 16, 325–332.
- Patuzzi, R.B., Johnstone, B.M., Sellick, P.M., 1984. The alteration of the vibration of the basilar membrane produced by loud sound. *Hear. Res.* 13, 99–100.
- Patuzzi, R.B., Yates, G.K., Johnstone, B.M., 1989a. Changes in cochlear microphonic and neural sensitivity produced by acoustic trauma. *Hear. Res.* 39, 189–202.
- Patuzzi, R.B., Yates, G.K., Johnstone, B.M., 1989b. OHC receptor current and sensorineural hearing loss. *Hear. Res.* 42, 47–72.
- Popelka, G., Osterhammel, P., Nielsen, L., Rasmussen, A., 1993. Growth of distortion product otoacoustic emissions with primary tone level in humans. *Hear. Res.* 71, 12–22.
- Robles, L., Ruggero, M.A., Rich, N.C., 1986. Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and phase responses. *J. Acoust. Soc. Am.* 80, 1364–1374.
- Ruggero, M.A., Robles, L., Rich, N.C., Recio, A., 1992. Basilar membrane responses to two-tone and broadband stimuli. *Phil. Trans. R. Soc. Lond.* 336, 307–315.
- Ruggero, M.A., Rich, N.C., Recio, A., Narayan, S.S., Robles, L., 1997. Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J. Acoust. Soc. Am.* 101, 2151–2163.
- Schmiedt, R.A., 1982. Boundaries of two-tone rate suppression of cochlear nerve activity. *Hear. Res.* 7, 335–351.
- Stover, L., Gorga, M.P., Neely, S.T., Montoya, D., 1996. Toward optimizing the clinical utility of distortion product otoacoustic emission measurements. *J. Acoust. Soc. Am.* 100, 956–967.
- Sun, X., Schmiedt, R.A., He, N., Lam, C.F., 1994a. Modelling the fine structure of the $2f_1 - f_2$ acoustic distortion product. I. Model development. *J. Acoust. Soc. Am.* 96, 2166–2174.
- Sun, X., Schmiedt, R.A., He, N., Lam, C.F., 1994b. Modelling the fine structure of the $2f_1 - f_2$ acoustic distortion product. II. Model evaluation. *J. Acoust. Soc. Am.* 96, 2175–2183.
- Viemeister, N.F., Bacon, S.P., 1988. Intensity discrimination, increment detection, and magnitude estimation for 1 kHz tones. *J. Acoust. Soc. Am.* 84, 172–178.
- Whitehead, M.L., Lonsbury-Martin, B.L., Martin, G.K., 1990. Actively and passively generated acoustic distortion at $2f_1 - f_2$ in rabbits. In: Dallos, P., Geisler, C., Matthews, J., Ruggero, M., Steele, C. (Eds.), *The Mechanics and Biophysics of Hearing*. Springer-Verlag, Madison, WI, pp. 243–250.
- Whitehead, M.L., Lonsbury-Martin, B.L., Martin, G.K., 1992a. Evidence for two discrete sources of $2f_1 - f_2$ distortion-product otoacoustic emissions in rabbit. I: Differential dependence on stimulus parameters. *J. Acoust. Soc. Am.* 91, 1587–1607.
- Whitehead, M.L., Lonsbury-Martin, B.L., Martin, G.K., 1992b. Evidence for two discrete sources of $2f_1 - f_2$ distortion-product otoacoustic emissions in rabbit. II: Differential physiological vulnerability. *J. Acoust. Soc. Am.* 92, 2662–2682.
- Yates, G.K., 1990. Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions. *Hear. Res.* 50, 145–162.
- Yates, G.K., 1991. Auditory-nerve spontaneous rates vary predictably with threshold. *Hear. Res.* 57, 57–62.
- Yates, G.K., Kirk, D.L., 1997. Electrically evoked otoacoustic emissions – Implications for the reverse transduction process. In: Palmer, A., Rees, A., Summerfield, A., Meddis, R. (Eds.), *Psychophysical and Physiological Advances in Hearing*. Whurr, London, pp. 39–45.
- Yates, G.K., Winter, I., Robertson, D., 1990. Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range. *Hear. Res.* 45, 203–220.
- Zwicker, E., 1980. Nonmonotonic behaviour of ($2f_1 - f_2$) explained by a saturation feedback model. *Hear. Res.* 2, 513–518.