Research Paper

Conversion of amplitude modulation to phase modulation in the human cochlea

Sho Otsuka\textsuperscript{a,b}, Shigeto Furukawa\textsuperscript{b}

\textsuperscript{a} Center for Frontier Medical Engineering, Chiba Univ, 1-33 Yayoicho, Inageku, Chiba 263-8522, Japan
\textsuperscript{b} NTT Communication Science Laboratories, NTT Corporation, 3-1, Morinosato-Wakamiya, Atsugi-shi, 243-8522, Japan

\textbf{A R T I C L E   I N F O}

Article history:
Received 14 April 2020
Revised 26 April 2021
Accepted 12 May 2021
Available online 16 June 2021

Keywords:
Amplitude modulation
Phase modulation
Otoacoustic emissions
Nonlinearity
Cochlea

\textbf{A B S T R A C T}

When an amplitude modulated signal with a constant-frequency carrier is fed into a generic nonlinear amplifier, the phase of the carrier of the output signal is also modulated. This phenomenon is referred to as amplitude-modulation-to-phase-modulation (AM-to-PM) conversion and regarded as an unwanted signal distortion in the field of electro-communication engineering. Herein, we offer evidence that AM-to-PM conversion also occurs in the human cochlea and that listeners can use the PM information effectively to process the AM of sounds. We recorded otoacoustic emissions (OAEs) evoked by AM signals. The results showed that the OAE phase was modulated at the same rate as the stimulus modulation. The magnitude of the AM-induced PM of the OAE peaked generally around the stimulus level corresponding to the compression point of individual cochlear input-output functions, as estimated using a psychoacoustic method. A computational cochlear model incorporating a nonlinear active process replicates the above-mentioned key features of the AM-induced PM observed in OAEs. These results indicate that AM-induced PM occurring at the cochlear partition can be estimated by measuring OAEs. Psychophysical experiments further revealed that, for individuals with higher sensitivity to PM, the PM magnitude is correlated with AM-detection performance. This result implies that the AM-induced PM information cannot be a dominant cue for AM detection, but listeners with higher sensitivity may partly rely on the AM-induced PM cue.

© 2021 The Author(s). Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

Evidence suggests that the frequency modulation (FM; or equivalently phase modulation, PM) of a sound wave, even when its envelope is flat, causes AM of the cochlear partition vibration at a particular location (Ghitza, 2001; Moore and Sek, 1996; Saberi and Hafter, 1995). This FM-to-AM conversion is not surprising when we consider that each location of the cochlear partition is tuned to a particular frequency. The amplitude of the cochlear partition vibration at a certain location increases (decreases) as the instantaneous frequency of the signal approaches (departs from) the optimal frequency for the location (Ghitza, 2001; Moore and Sek, 1996; Saberi and Hafter, 1995). It has been further shown that the FM-induced AM can be used as a cue for processing the FM of the sound (Moore and Sek, 1996; Saberi and Hafter, 1995).

The present study offers evidence that conversion also occurs in the opposite direction (AM to PM) in humans and that it affects AM perception. We conceived this idea upon considering the cochlear partition as a nonlinear transmission device that is, in fact, equipped with an active amplification mechanism (Hudspeth, 2008). It is known that when an AM signal with a constant-frequency carrier is fed into a generic nonlinear amplifier, the phase of the carrier of the output signal is also modulated (Whitaker, 2005). Physiological studies using static sinusoidal signals have indicated that the phase response of the cochlear partition varies with stimulus intensity when the stimulus frequency is close to its characteristic frequency (CF) and the concomitant amplitude response increases nonlinearly (Nuttall and Dolan, 1993; Ruggero et al., 1997). The level-dependent phase changes would underlie the cochlear AM-to-PM conversion, referred to the nonlinear mechanism.

Abbreviations: AM, amplitude modulation; PM, phase modulation; FM, frequency modulation; AM-to-PM, amplitude modulation to phase modulation; FM-to-AM, frequency modulation to amplitude modulation; BM, basilar membrane; CF, characteristic frequency; OAE, otoacoustic emission; AMOAE, amplitude-modulated tone evoked OAEs; AMDL, amplitude modulation detection limen; IDL, intensity discrimination limen; AN, auditory nerve; 2I-2AFC, two-interval two-alternative forced-choice.

* Corresponding author at: Center for Frontier Medical Engineering, Chiba Univ., 1-33 Yayoicho, Inageku, Chiba 263-8522, Japan.
E-mail address: otsuka.s@chiba-u.jp (S. Otsuka).

https://doi.org/10.1016/j.heares.2021.108274
0378-5955/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)
By contrast, AM-to-PM conversion can apparently also occur in a linear system (i.e., a system without an active nonlinear gain process) when it incorporates a non-flat frequency response, which distorts the relationship between phase and amplitude among spectral components of the AM signal (Whitaker, 2005). Hence the characteristics of cochlear AM-to-PM conversion cannot be predicted solely by the characteristics of the level-dependent phase response, namely the nonlinear versus linear mechanism.

In the present study, we first evaluated the possibility that AM-to-PM conversion occurs also in the otoacoustic emissions (OAEs), which are generated in the cochlear partition. OAEs have been used to probe the properties of cochlear partition motion around the peak of a traveling wave, such as delay (Shera et al., 2002) and frequency tuning (Keefe et al., 2008). If AM-to-PM conversion occurs in the cochlear partition, the carrier phase of OAEs should be modulated. Secondly, we examined whether AM-induced-PM information has appreciable impacts on the perception of AM by conducting psychoacoustic experiments. Lastly, the validity of the estimated characteristics of the AM-to-PM conversion was tested using a cochlear model incorporating a nonlinear active process.

2. Material and methods

2.1. Participants

44 volunteers (3 males and 41 females) aged 21–43 years (mean = 37.4, standard deviation = 5.9) participated in the study. All had normal pure-tone audiometric thresholds (HL ≤ 25 dB) ranging from 0.25 to 8 kHz. All participants were informed of the aims and risks of the experiment and gave written informed consent. The experiments were performed in accordance with the declaration of Helsinki of 1964 and approved by the Research Ethics Committee of NIT Communication Science Laboratories.

2.2. Equipment

Stimuli were digitally synthesized with sampling rates of 96 kHz and converted into analog signals using an audio interface (Edirol UA-101 and MOTU MicroBookltc for AMEOAE and psychoacoustic measurements, respectively) with a precision of 24 bits. For psychoacoustic measurements, the converted signals were presented through Sennheiser HDA 200 headphones. For AMEOAE measurements, the converted signals were amplified using a head- phone buffer and presented through Etymotic Research ER-2A earphones connected to an ER-10B low-noise microphone system. The stimulus was presented to the participant’s right ear. The two outputs from the ER-2A were calibrated using a Bruel and Kjaer Type 4257 ear simulator (IEC 711). The ear-canal sound pressure was recorded using an ER-10B inserted into each ear. All measurements were conducted in a double-wall sound-attenuating room.

2.3. OAE measurements

As a non-invasive measure of cochlear partition motion, we recorded OAEs evoked by an amplitude-modulated tone (AMEOAE: amplitude-modulated tone evoked OAEs) (Goodman, 2004) and examined their phase responses. AM tones were characterized using the following equation:

\[ s(t) = [1 + m \sin(2\pi f_a t)] \sin(2\pi f_c t) \]  

(1)

where \( f_c \) is the carrier frequency, \( m \) is the modulation index, and \( f_m \) is the modulation rate. The modulation depth was defined as \( \text{20log}_{10}(m) \). The \( f_c \) was set at 1 kHz or 4 kHz, and \( f_m \) was 50 Hz. The stimulus duration was 80 ms and included 10-ms raised-cosine ramps. The evoked AMEOAEs were measured using a microphone set in the ear canal. The recorded signals were mixtures of the AMEOAE and stimulus ringing in the ear canal. To eliminate the ringing, the nonlinear portion of the AMEOAEs was extracted using the double-evoked procedure (Keefe, 1998), where three AM tones defined by Eq. (1), \( S_1 \), \( S_2 \), and \( S_{12} \) (\( S_1 \) and \( S_2 \) were identical AM tones, and \( S_{12} \) means that \( S_1 \) and \( S_2 \) were presented at the same time) are presented separately, and the residual emission was extracted by the equation of \( S_1 + S_2 - S_{12} \). The stimulus sequence was presented 500 times for each condition. The average waveforms were band-pass filtered using a second-order Butterworth filter (925–1075 Hz). The envelope and phase of the average AMEOAEs were calculated using the Hilbert transform. Due to the presence of onset and offset ramps, only the middle portion (20–60 ms after stimulus onset) of the AMEOAE was analyzed. To thoroughly evaluate the properties of the cochlear AM-to-PM conversion, the dependency of AM-induced PM on the modulation depth and stimulus intensity was measured. The level dependence of the AMEOAE was measured by varying the stimulus level from 40 to 80 dB SPL in 5-dB steps while maintaining the modulation depth at ~7.2 dB, which is the dB relative to the primary level. The modulation depth dependence of the AMEOAE was measured by varying the modulation depth from -25 to -5 dB in 5-dB steps for stimulus levels of 60, 70, and 80 dB SPL. The modulation depth dependence of the AMEOAEs at 1 kHz was measured for thirteen female volunteers. The level dependences of the AMEOAEs at 1 and 4 kHz were measured for 44 volunteers (3 males and 41 females) and 19 volunteers (1 male and 18 females), respectively.

2.4. Psychoacoustic input-output function measurement

To explore how the AM-induced PM observed in the OAEs was linked to the input–output characteristics of the cochlear partition, we compared the magnitude of the AM-induced PM in the AMEOAEs as a function of stimulus level with the input–output function of the cochlear partition, which was estimated using a psychoacoustic method (Nelson et al., 2001). The temporal masking curve (TMC) method (Nelson et al., 2001) was used. The method measures forward masking threshold as a function of interval between the masker and the probe, referred to TMC. Assuming that forward-masking recovery is independent of masker level and probe frequency, the cochlear compression can be obtained by comparing thresholds for on- and off-frequency maskers, because threshold for the off-frequency masker can be considered as linear reference. TMCs were measured for a probe frequency of 1 kHz and for masker frequencies of 1 kHz (on-frequency) and 0.5 kHz (off-frequency). A two-interval two-alternative forced-choice (2I-2AFC) procedure and a two-down, one-up transformed adaptive method were used to track 70.7% (37) correct masked thresholds. The two intervals were presented to the listeners in random order, and the listeners were instructed to select the interval containing the probe. The inter-stimulus interval was 500 ms. The probe was kept at a sensation level of 9 dB. For both masker frequencies, the masker thresholds were measured in 5-ms steps of for mask-probe intervals ranging from 5 to 100 ms. The sinusoidal maskers were gated with 5-ms raised-cosine onset and offset ramps and had a total duration of 110 ms. The sinusoidal probes had a total duration of 10 ms and were gated with 5-ms raised-cosine ramps (no steady-state portion). One track of measurements was terminated after 12 reversals, and the masked threshold (in dB SPL) for that track was defined as the geometric mean of all the masker levels of the last eight reversals. The masker level step size was a factor of 6 dB for the first four reversals and a factor of 3 dB for later reversals. Each masked threshold was estimated as the geometric mean across three tracks. If the difference between masked thresholds exceeded 3 dB, two additional tracks were obtained. Input-output functions of the cochlear partition were inferred with the standard method by plotting the masker levels as a linear reference.
TMC against the masker levels for any other TMC, paired according to the masker–probe time interval (Nelson et al., 2001). The TMCs were measured for seven female volunteers.

2.5. AM detection limen measurement

To investigate whether the AM-to-PM conversion taking place in the cochlea has a significant impact on perception, we sought to determine whether the presence of AM-induced-PM information, in addition to the simple AM information, affects the AM detection performance of the human listener. We compared AM detection limens (AMDLs) measured psychophysi- cally and the PM magnitude measured in OAEs. The association between AMDLs and the AM-induced PM magnitude should be observed only when PM information coded by auditory nerve (AN) phase-locking was available. We confirmed this by repeating the same experiment, but this time adopting a carrier frequency of 4 kHz, rather than 1 kHz. AM detection limen was defined as the modulation depth at the detection threshold and was measured using a 2I-2AFC procedure and a two-down, one-up transformed adaptive method. Amplitude was modulated sinusoidally with a modulation depth of \( \Delta \delta \). The carrier frequency was 1 or 4 kHz, and the modulation frequency was 2 Hz. The stimulus duration was 750 ms, including 20-ms raised-cosine ramps. There was a 500-ms interval between the intervals. The starting modulator phase was randomly changed for each presentation. A run was started with \( \Delta \delta = 0.5 \). \( \Delta \delta \) was changed by a factor of 0.5 for the first four reversals and by a factor of 0.79 for later reversals. The AMDL for that track was defined as the geometric mean of all the \( \Delta \delta \) values of the last eight reversals. The mean AMDL for each participant was computed from two or four thresholds. The level dependence of the AMDL was measured by varying the stimulus level from 40 to 80 dB SPL in 5-dB steps. The AMDL at 1 and 4 kHz were measured for 44 volunteers (3 males and 41 females) and 19 volunteers (1 male and 18 females), respectively.

2.6. Intensity discrimination limen measurement

To investigate whether the observed association with the AM-induced PM magnitude were specific to amplitude modulation and do not reflect the variation in sensitivity to instantaneous amplitude or stimulus intensity. In another set of experiments, intensity discrimination limens (IDLs) were measured instead of AMDLs (the stimulus frequency was 1 kHz). The IDLs were measured using a 2I-2AFC procedure and a two-down, one-up transformed adaptive method. The “signal” and “non-signal” intervals in the 2I-2AFC method contained RSR and RRR sequences, respectively, where R indicates a 1-kHz pure tone as the reference and S indicates a 1-kHz pure tone with a \( \Delta \delta \) [dB] increment in level. Listeners were required to indicate the signal interval (RSR). \( \Delta \delta \) was changed in 3-dB steps for the first four reversals and in 1.5-dB steps for later reversals. The durations of R and S were 250 ms, including 20-ms raised-cosine ramps. There was no gap between the offset of the preceding tone and the onset of the following tone within the RSR or RRR interval. There was 500-ms silent gap between the signal and non-signal intervals. The phase of each tone was randomized every presentation to prevent listeners from using the level dependent phase changes to detect level changes. The level dependence of the IDL was measured by varying the stimulus level from 40 to 80 dB SPL in 5-dB steps. The level dependence of the IDL was measured for 35 volunteers (3 male and 32 females).

2.7. PM detection limen measurement

The ability of individual listeners to make use of the PM information generated by AM was evaluated by measuring PM detection limens (PMDLs). The PMDL was defined as the minimum detectable phase excursion of a sinusoidal signal and was measured using a 2I-2AFC procedure and a two-down, one-up transformed adaptive method. Phase was modulated sinusoidally with a phase excursion of \( \Delta \theta \). The carrier frequency and the modulation frequency were 1 kHz and 2 Hz, respectively. The stimulus duration was 750 ms, including 20-ms raised-cosine ramps. There was a 500-ms silent gap between the intervals. The starting \( \Delta \theta \) was set at 2.0 radian. \( \Delta \theta \) was changed by a factor of 0.5 for the first four reversals and by a factor of 0.79 for later reversals. The starting modulator phase was randomly changed for each presentation. Subsequent average procedures were the same as for the AMDL measurement. The stimuli were presented to the participant’s right ear at 50 dB SPL. The PMDL was measured for seventeen volun-
 teers (1 male and 16 females).

3. Results

3.1. Characteristics of AM-induced PM observed in OAEs

Consistently among the participants, the amplitude and phase of the AME0AEs were modulated in anti-phase with each other, and the AM-induced PM magnitude increased monotonically with modulation depth (a typical response is shown in Fig. 1A). The pattern was consistent across various sound pressure levels (50, 70 and 80 dB), and the slopes of the functions were significantly positive (Fig. 1B; \( T_{12} = 5.2, p = 0.021 \) for 60 dB; \( T_{12} = 8.2, p = 0.0022 \) for 70 dB; \( T_{12} = 5.1, p = 0.024 \) for 80 dB; Student’s t-test).

The AM-induced PM observed in OAEs was closely linked to an estimate of the cochlear partition’s input-output characteristics. Fig. 1C shows the AM-induced PM magnitude in the AME0AEs as a function of the stimulus level and the cochlear partition’s input-output function estimated by using a psychoacoustic method (Nelson et al., 2001) with independent experiments. We fitted a broken-stick function, defined by Eq. (2), to the individual psychoacoustic input-output (I/O) function using the least-squares method.

\[
\begin{align*}
\text{y} &= x + y_0 \ (x \leq x_C) \\
\text{y} &= A(x - x_C) \ (x > x_C)
\end{align*}
\]  

where \( x_C \) is the level at the compression point, \( A \) is the slope of the compressive part of the psychoacoustic I/O function, and \( y_0 \) is the intercept (the value of \( y \) when \( x = 0 \)). The mean of the estimation root mean square error (RMSE) when using the broken-stick function was 2.4 dB (standard deviation (SD) of 0.69 dB).

We found that the PM peak tended to appear slightly above the compression point. To quantify this tendency, we calculated the x-dB compression point of the psychoacoustic I/O function, which was defined as the input level at which a response decreased by x dB compared with a linear extrapolated line fitted to the lower level of the fitted broken-stick function (Fig. 1C). We found that the 3.6-dB compression point of the psychoacoustic I/O function was the optimum predictor for the variations, i.e., RMS difference between the x-dB compression point and the level with the PM peak was minimum at \( x = 3.6 \) (RMSE=7.7 dB; Pearson \( R_0=0.82, p=0.025 \); Figs. 1D and 1E).

3.2. Effect of the AM-induced PM on AM perception

Does the AM-to-PM conversion taking place in the cochlea have significant impacts on perception? As stated earlier, PM magnitude induced by a fixed modulation depth varies with stimulus level. We expected that the AMDL would vary with stimulus level in parallel with the PM magnitude. The carrier frequency in both measures was set at 1 kHz, where AN discharges can be phase locked to the temporal fine structure (TFS) of the cochlear partition vi-
participants’ 3.6-dB performance (temporal gestation frequency bration) (i.e., the AMDL is lower) at levels for which the PM was larger (Fig. 2A right). The former and latter tendencies, respectively, can be expressed as negative and positive correlation coefficients between the AMDL-versus-level function and the PM-magnitude-versus-level function ($R_{PM-AMDL}$; Fig. 2B). A Gaussian finite mixture model fitted to the $R_{PM-AMDL}$ distribution supported the notion of two components in the distribution (mean = -0.4; 95% confidence = -0.47 - -0.33 < 0 and mean = 0.27; 95% confidence = 0.19 – 0.34 > 0; 95% confidence was computed by a bootstrapping procedure). The likelihood of the two-Gaussian model

**Fig. 1.** Characteristics of AM-to-PM conversion observed in AMEOAEs. A: An example of amplitude and phase (relative to the stimulus phase) of AMEOAEs with different modulation depths of −25 to −5 dB (upper and lower panels, respectively). The responses were measured by using a sinusoidally amplitude-modulated tone with a carrier frequency of 1 kHz and a modulation frequency of 50 Hz. As with the simulated cochlear partition motion (Fig. 4C), the AMEOAEs exhibited increasing amplitude and phase modulation depth with increasing stimulus modulation depth. B: Relation between stimulus modulation depth and the PM magnitude observed in AMEOAEs for different stimulus levels (60, 70 and 80 dB SPL). The PM magnitude was defined as the standard deviation of the relative phase. The error bars show the standard errors of thirteen participants’ data. C: An example of a cochlear input-output (I/O) function estimated with a psychoacoustic method (upper panel) and the magnitude of PM observed in AMEOAEs as a function of stimulus level at 1 kHz (bottom panel). The solid line in the upper panel represents a broken-stick line fitted to the data points. The PM magnitude exhibited a peak around the level at which the gain of the I/O function changed. D: Root mean square (RMS) difference between x-dB compression point and level with PM peak ($L_{peak}$) as a function of x. x-dB compression point of psychoacoustic I/O function (arrow in C) was defined as input level at which response decreased by x-dB compared with linear extrapolated line (dotted line in C) fitted to lower part of fitted broken-stick function (solid line in C). E: Correspondence between $L_{peak}$ and the 3.6-dB compression point of the psychoacoustic I/O function, which was the best predictor of $L_{peak}$. Each dot represents one participant. $R$ depicts Pearson's correlation coefficient.
was significantly larger than that of the single-Gaussian model's ($p < 0.001$).

The diversity cannot be explained by the degree of conversion from AM to PM. There was no significant correlation between the average magnitude of the AM-induced PM across the levels and $R_{\text{PM-AMDL}}$ ($R_{41} = 0.12, p = 0.23$). An alternative and more likely contributory factor is the inter-individual variations of efficiency of the PM processing. To test this hypothesis, we measured the participant's sensitivity to PM imposed to tone bursts with a carrier frequency of 1 kHz and a modulation frequency of 2 Hz. As expected, listeners with a negative $R_{\text{PM-AMDL}}$ ($N = 7$) exhibited better sensitivity to PM than those with a positive $R_{\text{PM-AMDL}}$ ($N = 10$) ($T_{15} = 3.3, p = 0.0044$; Fig. 2C), i.e., listeners with a better PM sensitivity exhibited a tendency for the AM detection performance to be better at a level where the PM was larger.

The association between AMDL and the AM-induced PM magnitude should be observed only when PM information coded by AN phase-locking was available. We confirmed this by repeating the same experiment, but this time adopting a carrier frequency of 4 kHz, rather than 1 kHz. For that high frequency, AN phase locking is diminished or weakened (Johnson, 1980). As expected, the results showed no evidence of an association between PM magnitude and AMDL; a Gaussian finite mixture model fitted to the distribution of the correlation coefficients indicated that there was one component (average = -0.09; SD = 0.25), and the coefficient was not significantly different from zero ($T_{16} = -1.64; p = 0.12$).

We also confirmed that the observed association with the AM-induced PM magnitude were specific to amplitude modulation and do not reflect the variation in sensitivity to instantaneous amplitude or stimulus intensity. In the intensity discrimination task, listeners were required to detect a slight difference in level between two flat-envelope tone bursts, which were presented discontinuously. Since the starting phases of the tone bursts were randomized among the presentations, the phase information will have provided no clues for the task. Again, we did not find any associations between the PM magnitude and the IDLs; a Gaussian finite mixture model fitted to the distribution of the correlation coefficients showed that there was one component (average = 0.04; SD = 0.32), and there was no significant difference from zero ($T_{27} = 0.87; p = 0.39$).

### 3.3. Model analysis

To explore whether and how the AM-induced PM observed in OAEs reflects that observed at the cochlear partition, we compared the characteristics of the AM-to-PM conversion with those simulated by a computer model of the cochlea. A one-dimensional transmission line model, in which each portion of the cochlear partition is modeled as a mechanical element (i.e., masses, dampers, and springs), was constructed to capture the compressive nonlinear features, e.g., the level-dependent phase reported in physiological studies (Nuttall and Dolan, 1993; Ruggero et al., 1997). An
apparent AM-to-PM conversion can also occur in a linear system when it incorporates a non-flat frequency response, which distorts the amplitude and phase relation of the two sidebands of the AM signal (Whitaker, 2005). To determine the mechanisms underlying the AM-induced PM observed in the nonlinear cochlear model and AMEOAEs, we also simulated the PM of the cochlear partition motion induced via a linear mechanism, by using a quasi-linear cochlear model.

3.4. Cochlear model

The structure of the one-dimensional transmission line cochlear model was identical to that proposed by Ku and coworkers (Ku et al., 2009), except for the nonlinear characteristics implemented in the active process. A two-state Boltzmann function was used. The function of the nonlinear characteristics was simpler than the asymmetrical function used in Ku et al. (2009), but they are functionally the same. The cochlear partition was divided into 500 (N) segments along the length of the cochlea, I; the segment at the apical and basal end were models of the middle ear and the helicotrema, respectively. Those segments comprise a mass, spring and damper, which determine the boundary conditions. The other elements were a micromechanical model of the cochlear partition, which comprises two sets of masses, springs and dampers and mimics the structure of the cochlea (Fig. 3).

The cochlear micromechanics of the nth partition (n = from 2 to N-1) segments are described by the following equation:

$$\dot{x}_n(t) = A_n x_n(t) + B_n p_n(t)$$

(3)

where $x_n(t)$ is the vector of the velocity and the basilar membrane (BM) and tectorial membrane (TM) displacements at the nth micromechanical element and is equal to $[x_1[n] x_1[n] x_2[n] x_2[n]]^T$ (subscripts 1 and 2 indicate BM and TM, respectively; Fig. 3). $p_n(t)$ is the fluid pressure difference at the nth micromechanical model. $A_n$ and $B_n$ are defined as follows:

$$A_n = \begin{bmatrix}
-C_1[n] & 0 & 0 & 0 \\
-C_2[n] & -k_1[n] & 0 & 0 \\
-C_3[n] & 0 & -k_2[n] & 0 \\
-C_4[n] & 0 & 0 & -k_3[n]
\end{bmatrix}_{M 	imes M}$$

(4)

$$B_n = \begin{bmatrix}
\frac{1}{\rho} \\
0 \\
0 \\
0
\end{bmatrix}_{1 \times M}$$

(5)

$$x_1(t) = A_1 x_1(t) + B_1 p_1(t)$$

(8)

$$x_1(t) = \begin{bmatrix}
\dot{x}_m \\
x_m
\end{bmatrix}^T$$

(9)

$$A_1 = \begin{bmatrix}
-C_m & -k_m \\
0 & 0
\end{bmatrix}$$

(10)

The parameters, such as stiffness ($K_{1-4}[n]$) and damping ($C_{1-4}[n]$), varied continuously with the position on the BM and were identical to those employed by Ku et al. (2009) (Table 1). $\gamma_n$ is the nonlinear gain factor, which decreases with an increase in the driving level, and it is defined as

$$\gamma_n = \frac{1}{\beta \left(1 + e^{\frac{u_0}{\beta \gamma_n (x_0[n]-x_1[n])}}\right)}$$

(6)

where $u_0$ and $\beta$ represent the saturation point and the slope of the nonlinear function and are set at 0.5 and 10−12[m/s], respectively. $x_0[n]$ is the instantaneous shear displacement between the BM and the TM, i.e., $x_2[n] - x_1[n]$. In the quasi-linear cochlear model, different gain factors were set for each stimulus level to equate the velocity at the peak of the traveling wave to that simulated by the nonlinear model. Therefore, $\gamma_n$ was constant (0.2–1.0) for each stimulus level and did not depend on $x_0$. The active pressure ($P_{ac}$) can be determined using the following equation:

The middle ear dynamics (n = 1) can be written as follows:

$$x_1(t) = A_1 x_1(t) + B_1 p_1(t)$$

(8)

$$x_1(t) = \begin{bmatrix}
\dot{x}_m \\
x_m
\end{bmatrix}^T$$

(9)

$$A_1 = \begin{bmatrix}
-C_m & -k_m \\
0 & 0
\end{bmatrix}$$

(10)

---

**Table 1** Cochlear model parameters (adapted from Ku et al. (2009)). $x$ is the longitudinal distance along the cochlea.

| Quality | Formula (SI) | Units |
|---------|--------------|-------|
| $k_1$ | $1.65 \times 10^3 e^{-7ln(x-0.00173)}$ | N m\(^{-1}\) |
| $C_1$ | $9 \times 9990e^{-153(x-0.00173)}$ | N m\(^{-1}\) |
| $m_1$ | $4.5 \times 10^{-3}$ | kg m\(^{-1}\) |
| $k_2$ | $1.05 \times 10^3 e^{-107(x-0.00173)}$ | N m\(^{-1}\) |
| $C_2$ | $300e^{-172(x-0.00173)}$ | N m\(^{-1}\) |
| $m_2$ | $7.20 \times 10^{-4} + 2.87 \times 10^{-2} x$ | kg m\(^{-1}\) |
| $k_3$ | $1.5 \times 10^3 e^{-275(x-0.00173)}$ | N m\(^{-1}\) |
| $C_3$ | $6.6e^{-55}(x-0.00173)$ | N m\(^{-1}\) |
| $k_4$ | $9.23 \times 10^3 e^{-107(x-0.00173)}$ | N m\(^{-1}\) |
| $C_4$ | $3300e^{-144(x-0.00173)}$ | N m\(^{-1}\) |
| $k_m$ | $2.63 \times 10^0$ | N m\(^{-1}\) |
| $C_m$ | $2.8 \times 10^6$ | N m\(^{-1}\) |
| $m_b$ | $2.96 \times 10^{-2}$ | kg m\(^{-1}\) |
| $c_b$ | $350$ | kg m\(^{-1}\) |
| $\rho$ | $1000$ | kg m\(^{-1}\) |
| $H$ | $0.001$ | m |
| $L$ | $0.035$ | m |
\[ B_1 = \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & 0 \end{bmatrix} \]  

(11)

The helicotrema dynamics \((n = N)\) can be written as follows:

\[ \dot{x}_N(t) = A_N x_N(t) + B_N p_N(t) \]  

(12)

\[ x_n(t) = \begin{bmatrix} \dot{x}_n \\ x_n \end{bmatrix} \]  

(13)

\[ A_N = \begin{bmatrix} m & 0 \\ 1 & 0 \end{bmatrix} \]  

(14)

\[ B_N = \begin{bmatrix} \frac{1}{m} \\ 0 \end{bmatrix} \]  

(15)

These cochlear segments were coupled with fluid mechanics using the one-dimensional wave equation, and the coupled relationship is described by the following state-space formulation:

\[ \dot{x}(t) = A x(t) + B u(t) \]  

(16)

where \(x(t)\) is the vector of state variables, which include the BM and TM velocities and displacements.

\[ x'(t) = \begin{bmatrix} x_1(t) & x_2(t) & \cdots & x_{N-1}(t) & x_N(t) \end{bmatrix}^T \]  

(17)

\[ A \] is the system matrix that describes the coupled mechanics, and \(B\) is the input matrix.

\[ A = \begin{bmatrix} 1 - B_1 F^{-1} C_e \end{bmatrix}^{-1} \]  

(18)

\[ B = \begin{bmatrix} 1 - B_1 F^{-1} C_e \end{bmatrix}^{-1} B_e \]  

(19)

where \(A_e\) and \(B_e\) are block diagonals defined by the following equations:

\[ A_e = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 \end{bmatrix} \]  

(20)

\[ B_e = \begin{bmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 \end{bmatrix} \]  

(21)

\(F\) is the fluid-coupling matrix.

\[ F = \frac{H}{2 \rho \Delta^2} \begin{bmatrix} -\frac{3}{\pi} & -\frac{3}{\pi} & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ -2 & -2 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{2 \rho \Delta^2}{H} \end{bmatrix} \]  

(22)

where \(\rho\) is the density of the cochlear fluids, and \(H\) is the height of the canal above and below the cochlear partition. The cochlear length \(L\) is divided into \(N\) (500) sections of length \(\Delta\).

\(u(t)\) is a vector of inputs equal to \(F^{-1} q(t)\), \(q(t)\) is the vector of source terms that serves as the input to the micromechanical model.

\[ q(t) = \begin{bmatrix} w_{50}(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \]  

(23)

where \(w_{50}(t)\) is the acceleration due to the pressure in the ear canal. \(w_{50}(t)\) was calculated from the ear canal pressure using a two-port network model of the middle ear (Kringlebotn, 1988). The differential equation was solved using the Runge-Kutta 4th-order algorithm. The step size for the algorithm was set to \(3 \times 10^{-6}\) s.

### 3.5. AM-induced PM simulated by nonlinear cochlear model

The intensity dependence of the simulated phase response of the BM is shown in Fig. 4A, B. At basal and apical locations, the phase increases or decreases as level increases, respectively, which is qualitatively similar to the data reported in previous animal studies (Nuttall and Dolan, 1993; Ruggero et al., 1997). Fig. 4C is a plot of the motion of a simulated BM, at the position of a traveling-wave peak and at positions slightly basal and apical with respect to that (0.56 mm from the peak position), driven by sinusoidally amplitude-modulated tones with modulation depths in the range of -25 to -5 dB. Similar to the OAE results (Fig. 1A, B), at the traveling-wave peak and the slightly basal position in the BM, both the envelope (top row) and phase (bottom row) of the BM motion are modulated in anti-phase; in addition, the magnitude of the AM-induced PM increases monotonically with the modulation depth. However, at the slightly apical position with respect to the peak, the envelope and phase are modulated in-phase.

The association between the magnitude of the AM-induced PM and the characteristics of the input–output function of the BM was also studied using the nonlinear cochlear model. Fig. 4D shows the input–output function of the simulated BM (top panel) and the magnitude of the AM-induced PM (middle panel) of the motion of a simulated BM at multiple positions around the traveling-wave peak. The magnitude of the AM-induced PM monotonically increases with distance from the traveling-wave peak (compare the peak heights in the middle panels, noting the difference in the ordinate scale). The pattern of the level-dependence of the PM, however, is clearly different for the basal and apical positions: near the peak and basal positions of the traveling-wave, as for the OAE results (Fig. 1C), the magnitude of the AM-induced PM tends to reach its maximum near to the compression point of the input-output function (compare top and middle panels in Fig. 4D). The level at which the AM-induced PM reaches maximum was estimated to be the -0.2-, 9.8-, and 10.3-dB compression points of the input-output function for 0.84-mm basal, 0.42-mm basal, and peak positions, respectively. At slightly apical positions with respect to the peak, in contrast, the magnitude of the PM induced by the nonlinear model is likely to reach a minimum near to the compression point of the input–output function.

Taken together, the key features of the AM-induced PM in OAEs are qualitatively similar to those observed in the simulated BM motion at the traveling-wave peak or at the slightly basal position.

### 3.6. AM-to-PM conversion predicted by BM level-dependent phase

The differences in the relative PM/AM phases between the basal and apical positions (Fig. 4C) can be rationalized by the level-dependent BM motion phase changes, i.e., the nonlinear mechanism (Nuttall and Dolan, 1993; Ruggero et al., 1997; Fig. 4A). At basal positions, the BM motion phase decreases as the level increases (Fig. 4A). The phase decrease associated with a level increase delays the phase near the maximum AM amplitude and advances the phase at the minimum AM amplitude (Fig. 5). Hence, the envelope and phase of the BM motion should be modulated anti-phase at the basal positions.

In contrast, at the apical positions, the BM motion phase increases as the level increases (Fig. 4A). The phase increase associated with level increase advances the phase near the maximum AM amplitude and delays the phase at the minimum AM amplitude (Fig. 5). Thus, the envelope and phase should be modulated in phase.

Regarding the basal positions, the level-dependence of the AM-induced PM (Fig. 4D) can also be explained by considering the level-dependent BM phase changes. Around the compression point of the input–output function of the BM, the BM phase change per
dB reaches its positive maximum at apical locations and its negative minimum at basal locations (bottom panels in Fig. 4B). Because a steeper slope of phase response, regardless to negative or positive, induces larger PM (Fig. 5), the magnitude of the AM-induced PM should be maximized close to the compression point of the BM input–output function, irrespective of the locations on the BM.

At the apical positions, however, the magnitude of the PM induced by the nonlinear model reaches its minimum near the compression point, and this cannot be explained by the level-dependent phase changes. This discrepancy suggests that the AM-induced PM observed at the apical locations is not generated solely by the nonlinear mechanism, i.e., it is not entirely a level-dependent characteristic.

**Fig. 4.** Effects of AM-to-PM conversion on simulated BM motion. A: Intensity dependence of simulated BM phase response around CF (≈1 kHz). BM phases are expressed relative to the phases of responses at 75 dB SPL. B: Velocity-intensity functions for simulated BM motion (top panel) and BM phase changes per dB (bottom panel) in response to tones with frequencies near CF. C: Velocities and phases (relative to the stimulus phase) of the simulated BM motion at positions near and at the traveling-wave peak generated by a sinusoidally amplitude-modulated tone (upper and lower panels, respectively). The stimulus level, carrier frequency, and modulation frequency were 60 dB SPL, 1 kHz, and 50 Hz, respectively. The depth of the amplitude and phase modulations of the BM motion increase as the stimulus modulation depth increases in the range from -25 to -5 dB. D: BM velocity and PM magnitude as a function of stimulus level at positions around the peak of the traveling wave (top and middle panels, respectively). The relative phase of the PM to with respect to the AM is plotted as a function of stimulus level in the bottom panel. Each response was generated by a sinusoidally amplitude-modulated tone with a modulation depth of -7.2 dB. The PM magnitude is defined as the standard deviation of the relative phase. Data simulated by the quasi-linear cochlear model is plotted as dashed lines in each panel. Different gain factors were set for each stimulus level so that the velocity at the peak of the traveling wave was equal to that simulated by the nonlinear model.
3.7. AM-induced PM simulated by a quasi-linear cochlear model

An AM-to-PM conversion was also observed in the quasi-linear cochlear model (i.e., a system without an active nonlinear gain process). The magnitude and phase of the PM produced by the quasi-linear cochlear model are plotted in Fig. 4D (middle and bottom panels, respectively). For the linear part of the input–output function, the phases and the magnitudes of the PMs generated using the quasi-linear and nonlinear cochlear models overlap. This implies that the PM, at low levels, is dominated by a linear mechanism, regardless of position on the BM. In contrast, near the compression point of the input–output function, the magnitude of the PM generated by the nonlinear cochlear model is markedly different from that generated by the quasi-linear cochlear model. These results imply that key features of the AM-induced PM observed in the nonlinear cochlear model and AMEOAEs cannot be accounted for by the linear mechanism.

4. Discussion

4.1. AM-to-PM conversion revealed by OAE measurements

We offer the first experimental evidence that AM-to-PM conversion also occurs in the human cochlea. We recorded OAEs evoked by an AM signal and showed that the phases of the OAEs were modulated at the same rate as the stimulus modulation. Although Goodman et al. (2004) measured OAEs evoked by AM signals and confirmed that the envelope of the OAEs was modulated at the same rate as the stimulus, they did not analyze the phase of the OAEs. In addition, we found that the magnitude of the AM-induced PM in the OAEs peaked typically near to the stimulus level at the compression point of individual cochlear input–output functions, as estimated by a psychoacoustic method. The relationship between the AM-induced PM and the psychoacoustic input–output function is qualitatively similar to that observed in the simulated BM motion at the traveling-wave peak or the slightly basal position. This is consistent with a widely-held assumption that OAEs are composed of reflected waves that originate near the peak and basal positions of the traveling-wave (Choi et al., 2008). Considering that the magnitude of the PM observed in the AMEOAE was larger than the PM at the traveling-wave peak in the cochlear model, we speculate that the AMEOAE PM is dominated by the PM generated at basal positions. The consistency between the OAEs results and those of the cochlear simulations confirms that the properties of AM-to-PM conversion at the cochlear partition can be evaluated, at least qualitatively, by measuring AMEOAEs.

4.2. Underlying mechanisms of AM-to-PM conversion

At the traveling-wave slightly basal positions, the level-dependence of the AM-induced PM at the simulated BM (Fig. 4D) can be rationalized by considering the characteristics of the level-dependent BM phase changes, i.e., the nonlinear mechanism. At the apical positions, however, the magnitude of the PM reaches its minimum close to the compression point, and this cannot be explained in terms of level-dependent phase changes. Because a steeper slope of phase response, regardless as to negative or positive, induces larger PM (Fig. 5), the magnitude of the PM induced by the nonlinear mechanism should be maximum close to the compression point of the BM input–output function, irrespective of the locations on the BM.

An AM-to-PM conversion was also observed in the quasi-linear cochlear model (i.e., a system without an active nonlinear gain process). This is consistent with the phenomena observed in an amplifier incorporating a non-flat frequency response, which distorts the amplitude and phase relation of two sidebands of an AM signal (Whitaker, 2005). The comparison between the nonlinear and quasi-linear model revealed that the dominant contribution to the
PM at low levels was the linear mechanism, regardless of the position on the BM (Fig. 4D). This is reasonable because the input-output function is almost linear, and the level-dependent phase change is very small at low levels (Fig. 6). In contrast, around the compression point of the input-output function the magnitude of the PM generated by the quasi-linear cochlear model was markedly different from that generated by the nonlinear cochlear model. The size of the AM-induced PM reaches minimum or maximum around the compression point in the nonlinear cochlear model, whereas the PM magnitude declined monotonically or almost flat across levels in the quasi-linear cochlear model. Although PM generated by the nonlinear and linear mechanism should interfere at middle to high levels, the pattern of AM-induced PM observed at basal locations, i.e., the maximum magnitude around compression point, is likely to be dominated by the nonlinear mechanism at middle to high levels. However, the pattern of the AM-induced PM observed at apical locations, i.e., the minimum magnitude around compression point, is not generated solely by the nonlinear or linear mechanism. Because the relative phase of the PM, with respect to that of the AM, was quite different between the quasi-linear and nonlinear cochlear model at apical positions (Fig. 4D), the PM induced by the nonlinear and linear mechanisms should interfere with each other. This would cause the discrepancy in level-dependence of AM-induced PM magnitude between the PM for the simulated BM and the prediction based on the level-dependent phase pattern.

Taken together, our results indicate that both the nonlinear and linear mechanisms contribute to the AM-to-PM conversion, but the common features of the AM-induced PM observed in the nonlinear cochlear model and AMEOAEs, i.e., maximum PM around compression points of input-output functions and anti-phase PM relative to AM, reflect the characteristics of the nonlinear mechanism.

4.3. Effect of the AM-induced PM on AM perception

Individuals with higher sensitivity to PM exhibited a tendency toward better AM detection performance (i.e., the AMDL is lower) at levels corresponding to higher PM. In addition, there is no relationship between the PM magnitude and performance in the intensity discrimination task - in which the phase information provides no clues for discrimination - and of AM detection at 4 kHz. For such a high frequency, AN phase locking is diminished or weakened (Johnson, 1980).

These results imply that the listeners with higher sensitivity to PM can make use of the phase-locked discharge pattern generated by AM-induced-PM on the cochlear partition for AM detection. This is consistent with our previous study (Otsuka et al., 2014), which demonstrated that low-rate (2 Hz) AMDL is correlated with the psychoacoustic indices of TFS sensitivity, e.g., low-rate (2 Hz) frequency-modulation detection limens and interaural phase difference thresholds (Otsuka et al., 2016, 2014; Strelcyk and Dau, 2009). Given relatively weak correlation between level dependent patterns of AM-detection performance and PM magnitude ($r = 0.47$), however, the AM-induced PM cue is not likely to be a dominant cue even for listeners with higher sensitivity to PM.

In contrast, individuals with lower sensitivity to PM exhibited a positive correlation between AMDL and the PM magnitude in OAE; i.e., the AM detection performance was worse when the PM was larger (Fig. 2C). A naive interpretation of this correlation is that AM-induced PM disrupts AM detection; however, we do not have a suggestion on the specific mechanism for this disruption. Another argument is that the level dependence of the AM-induced-PM amplitude (Fig. 4D; nonlinear model) exhibits an opposite trend in the apical and basal locations of the BM, and that the PM in OAE is considered to predominantly represent the AM-induced PM at basal locations. That is, at the apical locations, the magnitude of AM-induced PM is minimum near the compression point of the IOJ function, which is in contrast to the PM at the basal locations (see the middle row of panels in Fig. 4D). Therefore, if the apical PM information plays a dominant role in AM detection, AMDL would be the highest (i.e., poorest AM detection) near the compression point. The opposite trends observed in PM-sensitive and PM-insensitive individuals may be explained by their relative reliance on the basal and apical information, respectively, in the AM detection task. However, unfortunately, we do not have an explanation for the association between the individual PM sensitivity and the site dominance (i.e., apical versus basal BM).

Given that the AM-induced PM on the BM conveys somewhat useful information, how does the auditory system decode it? The computation of cross-correlations between the outputs from two distinct regions on the cochlear partition is a possible physiological mechanism. This type of mechanism has been suggested for
the perception of pitch, for detecting tones in noise, for encoding sound intensity, and for enhancing spectral representations of complex sounds (Carloy et al., 2012; Carney, 1994; Carney et al., 2002; Cedolin and Delgutte, 2010; Heinz et al., 2001; Loeb et al., 1983; Shamna, 1985).

Compared with the AM code, the PM code would be advantageous, for example, in processing middle and high levels of stimuli. The efficiency of the AM code based on changes in the AN discharge rate degrades severely for those levels because of the limited dynamic range of the AN discharge rate (May and Sachs, 1992). On the other hand, the phase-locking characteristics of the AN are less affected by stimulus intensity. Heinz and coworkers proposed a phenomenological model of intensity discrimination based on the level-dependent phase shift and showed that the model can explain the high performance intensity discrimination exhibited by human beings at high stimulus levels (Heinz et al., 2001). Nonlinear mechanical amplification is ubiquitous among species, such as mammalian, vertebrate etc., and therefore similar AM-to-PM conversion should occur across species. AM is contained in most natural sounds including vocalizations, and the PM code may also be adapted for the processing of ethological stimuli regardless of species.

**Funding**

This research was supported by internal research funding from NTT Corporation.

**Author contributions**

Sho Otsuka: Conceptualization, Data curation, Investigations. Formal analysis, Writing - original draft. Shigeto Furukawa: Conceptualization, Project administration, Methodology, Supervision and Writing - review and editing.

**Declaration of competing interests**

The authors declare no competing interests.

**References**

Carney, L.H., Heizny, M.G., Evislizer, M.E., Gilkeyz, R.H., Colburn, H.S., 2002. Auditory Phase opportunity: a temporal model for masked detection at low frequencies. Acta Acust. United Acust. 88, 334–347.

Cedolin, L., Delgutte, B., 2010. Spatio-temporal representation of the pitch of harmonic complex tones in the auditory nerve. J. Neurosci. Off. J. Soc. Neurosci. 30, 12712–12724. doi:10.1523/JNEUROSCI.6365-09.2010.

Choi, Y.-S., Lee, S.-Y., Parham, K., Neely, S.T., Kim, D.O., 2008. Stimulus-frequency otoacoustic emission: measurements in humans and simulations with an active cochlear model. J. Acoust. Soc. Am. 123, 2651–2669. doi:10.1121/1.2902184.

Chizita, O., 2001. On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception. J. Acoust. Soc. Am. 110, 1626–1640. doi:10.1121/1.1396325.

Goodman, S., 2004. Cochlear delays measured with amplitude-modulated tone-burst-evoked OAEs. Hear. Res. 188, 57–69. doi:10.1016/S0378-5955(03)00375-7.

Heinz, M.G., Colburn, H.S., Carney, L.H., 2001. Rate and timing cues associated with the cochlear amplifier: Level discrimination based on monaural cross-frequency coincidence detection. J. Acoust. Soc. Am. 110, 2065–2084. doi:10.1121/1.1404977.

Hudspeth, A.J., 2008. Making an effort to listen: mechanical amplification in the ear. Neuron 59, 536–545. doi:10.1016/j.neuron.2008.07.012.

Johnson, D.H., 1980. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J. Acoust. Soc. Am. 68, 1115–1122. doi:10.1121/1.384982.

Keefe, D.H., 1998. Double-evoked otoacoustic emissions. I. Measurement theory and nonlinear coherence. J. Acoust. Soc. Am. 103, 3489. doi:10.1121/1.423057.

Keefe, D.H., Ellison, J.C., Fitzpatrick, D.F., Gorga, M.P., 2008. Two-tone suppression of stimulus frequency otoacoustic emissions. J. Acoust. Soc. Am. 123, 1479–1494. doi:10.1121/1.2828209.

Kringelbom, M., 1988. Network model for the human middle ear. Scand. Audiol. 17, 75–85. doi:10.3109/01050398809070695.

Ku, E.M., Elliott, S.J., Lineton, B., 2009. Limit cycle oscillations in a nonlinear state space model of the human cochlea. J. Acoust. Soc. Am. 126, 739. doi:10.1121/1.3158861.

Loeb, G.E., White, M.W., Merzenich, M.M., 1983. Spatial cross-correlation. Biol. Cybern. 47, 149–163. doi:10.1007/BF00337005.

May, B.J., Sachs, M.B., 1992. Dynamic range of neural rate responses in the ventral cochlear nucleus of awake cats. J. Neurophysiol. 68, 1589–1602.

Moore, B.C.J., Sek, A., 1996. Detection of frequency modulation at low modulation rates: evidence for a mechanism based on phase locking. J. Acoust. Soc. Am. 100, 2320–2331. doi:10.1121/1.417941.

Nelson, D.A., Schroder, A.C., Wojtczak, M., 2001. A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 110, 2045–2064. doi:10.1121/1.1404439.

Nutall, A.L., Dolan, D.F., 1993. Two-tone suppression of inner hair cell and basilar membrane responses in the guinea pig. J. Acoust. Soc. Am. 93, 390–400. doi:10.1121/1.405619.

Ruggiero, M.A., Rich, N.C., Recio, A., Narayana, S.S., Robles, L., 1997. Basilar-membrane responses to tones at the base of the chinchilla cochlea. J. Acoust. Soc. Am. 101, 2151–2163. doi:10.1121/1.418265.

Saberi, K., Halfert, E.R., 1995. A common neural code for frequency- and amplitude-modulated sounds. Nature 374, 537–539. doi:10.1038/374537a0.

Shamna, S.A., 1985. Speech processing in the auditory system II: lateral inhibition and the central processing of speech evoked activity in the auditory nerve. J. Acoust. Soc. Am. 78, 1622–1632. doi:10.1121/1.392800.

Shera, C.A., Guinan, J.J., Osenbach, A.J., 2002. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc. Natl. Acad. Sci. 99, 3318–3323. doi:10.1073/pnas.032675099.

Whitaker, J.C., 2005. The Electronics Handbook, Second Edition CRC Press.