31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

High resolution system for improved transient-evoked otoacoustic emission acquisition

Christopher L. Bennett, Özcan Özdamar

Abstract—**Transient-evoked otoacoustic emissions (TEOAE) are generated by the cochlea in response to clicks. They are obtained by averaging post-onset acoustic responses which are composed of the stimulus-related meatal response (MR) and the TEOAEs. TEOAEs are typically below normal hearing thresholds and are obstructed by the MR, which is several orders of magnitude higher. For click stimuli, MRs typically last about 5 ms and obstruct the early latency emissions. TEOAEs become compressively nonlinear as the MR increases, and this property is commonly exploited by obtaining a derived nonlinear response (DNLR) which reduces the MR artifact. In this study we report the development of a high-resolution system which linearly acquires both MRs and TEOAEs. Results show that the duration of the artifact can be reduced, making the high frequency content of TEOAEs observable.**

I. Introduction

TRANSIENT-evoked otoacoustic emissions (TEOAE) are sounds that are generated within the cochlea in response to short-duration stimuli, such as clicks and tone pips [1]. These emissions are predicted in the active cochlear model of the ear, as outer hair cells enhance the backwards traveling coherent reflections caused by impedance variations along the basilar membrane [2], [3]. TEOAEs are extremely prevalent among the normal hearing population, and are commonly used as a hearing screening tool [4]. In the acquisition of TEOAEs, the stimulus and response are simultaneously recorded, with the stimulus lasting approximately 5 ms and the response lasting about 20 ms. The acoustic stimulus, sometimes called the meatal response (MR), is considered an artifact in the recording, as it obscures the early latency TEOAEs from being observed. This is due in part to the fact that the stimulus, depending on its intensity level, can be upwards of 50 dB higher in intensity than the response. Several methods have been proposed for the removal of the stimulus artifact from the response, including linear prediction [5], independent component analysis [6], wavelet based methods [7], digital subtraction [8], and the derived nonlinear response (DNLR) method [9]–[11], and a two-source DNLR method [12]. Among these, the DNLR method is the most commonly used; only the input stimuli are varied with little additional computation needed, which allows for simple integration into a linear-mode system.

In a DNLR recording, the peak TEOAE response, R_{max}, will typically be less than 1 mPa in amplitude. While the MR will vary, its peak amplitude, MR_{max}, can reach 250 mPa, as in the case of an 82 dB-SPL stimulus. The number of bits needed to cover this dynamic range is given by the $\log_2$ of the quotient of these numbers. The result must be rounded up (indicated by the ceiling function) since the number of bits are an element of the set of positive integers. In the case of a medium bit-depth system, the analog-to-digital converter (ADC) has 12 or 16 bit resolution, given by N_{system}. For these systems, it is not feasible to record the entire dynamic range of the MR and reserve a sufficient number of bits for the TEOAE, given by N_{TEOAE}. For example, if the entire amplitude-range of the MR is recorded via a 12-bit ADC (such as in the Otodynamics ILO88), then only 4 bits remain for the TEOAE, corresponding to a -24 dB digital noise floor (see Fig. 1a, Eqn. 1).

$$N_{system} = N_{TEOAE} + \left\lceil \log_2\left(\frac{MR_{max}}{R_{max}}\right)\right\rceil \tag{1}$$

On the other hand, if all 12 bits are reserved for the TEOAE, then a more appropriate digital noise floor of -72 dB can be obtained, however the stimulus will saturate the ADC (see Fig. 1b). In order to recover both the MR and the TEOAE, a higher bit-depth will be needed. In fact, if 12 bits are reserved exclusively for the TEOAE, which could have a

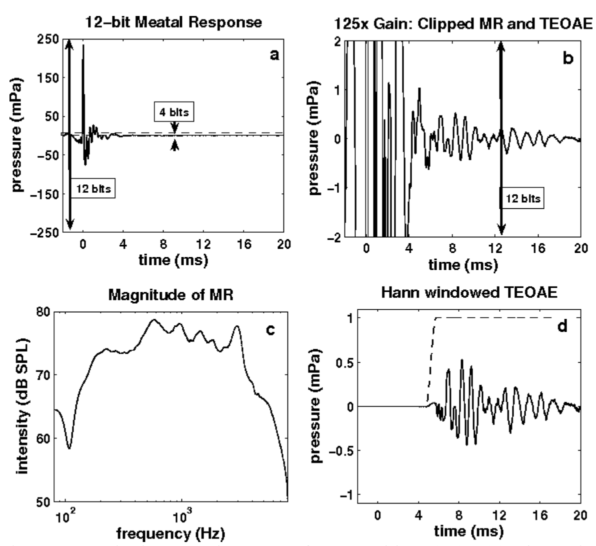

Fig. 1. A typical MR and TEOAE from a subject (01). (a) The entire metal response recovered using a 12-bit system leaves only 4 bits for representation of the TEOAE. (b) 125x gain applied to the recovered response which gives a full 12 bits to the TEOAE, but saturates the meatal response. (c) The spectrum of the meatal response. (d) 1 ms Hann window (of the raised Cosine family) applied at 5 ms to remove the meatal artifact.

Manuscript received April 6, 2009.

C.L. Bennett is with the Biomedical Engineering Department, University of Miami, Coral Gables, FL 33146 USA (phone: 305-284-5272; fax: 303-284-6494; e-mail: c.bennett@umiami.edu).

Ö. Özdamar is with the Department of Biomedical Engineering at the University of Miami, and also with the Departments of Otolaryngology, Pediatrics, and Neuroscience at the University of Miami Miller School of Medicine, Miami, FL (e-mail: oozdamar@miami.edu).

magnitude of approximately 1 mPa (on the high end), then in order to also recover the meatal response, or stimulus artifact, then at least 20 bits will be required. If greater bit-depth than 12 is desired for the TEOAE, then the system bit-depth must be comparably increased as well (eg., a 24-bit ADC is required if reserving 16 bits for N_{TEOAE}).

The frequency response of the meatus is subject-dependent, but will generally have a resonant peak in the 3-4 kHz range due to the length and volume of the meatus (see Fig. 1c). The ringing of the MR obscures the TEOAE, so windowing is used to eliminate the artifact, and subsequently the early latency TEOAEs as well (see Fig. 1d).

It was shown in [13] that the response of TEOAEs begins to saturate with increasing stimulus intensity. The MR is approximately linear with respect to the stimulus level, however the TEOAE response is nonlinear and grows compressively. So at high stimulus levels, an increase in intensity will comparably increase the amplitude of the MR, but only marginally increase the amplitude of the TEOAE. The growth of the TEOAE with respect to the input stimulus level can be given by a dB/dB slope. A slope of 1 dB/dB represents perfect linearity, and a slope of 0 dB/dB represents complete saturation [14]. It was found in [15] that a slope of 0.33 dB/dB exists in the nonlinear region.

The DNLR recovery method is a technique to partially cancel the stimulus artifact from the recorded TEOAE signal. The DNLR method exploits the compressive nonlinearities of TEOAEs by averaging groups of four responses; three of which are in the linear range of TEOAE response, and the fourth which is opposite polarity, and three times the magnitude, and thus in the nonlinear range of TEOAE response. The stimulus artifact is assumed to be linear, whereas the TEOAE response is compressively nonlinear. By summing the four MRs, the stimulus will theoretically cancel itself out. However, in a situation where the MR is saturating the ADC, as in a 12 or 16-bit system, only part of the MR is entirely obtained. This portion will, for the most part cancel, but the portion that saturates the ADC will not be able to cancel (see Fig. 2). 24-bit acquisition of the MR allows for near-complete cancellation of the entire artifact. With improved recovery of TEOAEs, it becomes possible to implement linear operands for further post-processing or signal analysis.

II. Material and Methods

A. Subjects

Data were acquired from 10 ears of 5 volunteers (3 male and 2 female) of ages ranging from 22 to 31, all in accordance with an IRB-approved protocol. All subjects had click-stimulus thresholds better than 25 dB-HL. For TEOAE acquisition, the subjects sat or lied down in an acoustically attenuated environment. An Etymōtic Research ER-10D OAE probe fitted with a rubber tip was inserted into the meatus to form an acoustic seal with the meatal wall.

B. Equipment and signal processing

Fig. 3 shows the block diagram of the system designed. The probe interfaced with a digital signal processor (Analog

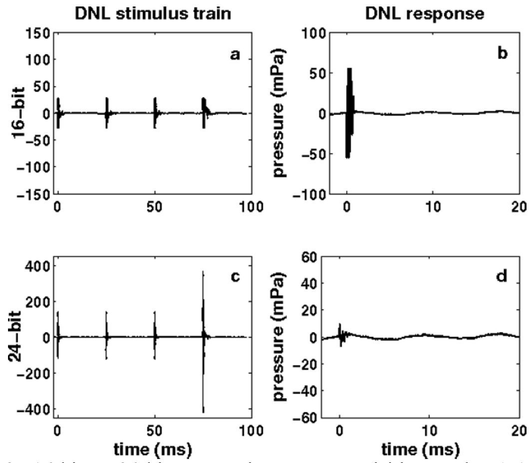

Fig. 2. 16-bit vs. 24-bit systems in DNLR acquisition mode. A 16-bit system, panels (a) and (b), will saturate the ADC and result in imperfect cancellation of the MR artifact in the 0-5 ms region. A 24-bit system, panels (c) and (d), can fully capture the MR and so artifact cancellation is much improved.

Devices ADSP-21364 Sharc EZ-Kit Lite) with an on-board ADC/DAC (AD1835: 24-bit, 48kHz). The output of the DAC was calibrated to the sound source specifications of the ER-10D probe, in which a 1 V_{RMS} signal is equivalent to 86 dB-SPL. Since the AD1835 DAC has a maximum peak output voltage of 1 V, then the maximum stimulus intensity level is 83 dB-SPL, or about 283 mPa. The ADC voltage levels were calibrated to pressure by taking into account the voltage range of the ADC, V_{pp}, and the sensitivity of the probe microphone, V_{sens}. The microphone sensitivity indicates how the microphone converts acoustic pressure into voltage and is given as 50 $^{mV}/_{Pa}$. The digital-to-pressure scaling equation is given in Eqn. 2.

$$R_{mPa} = \frac{V_{pp}}{2^{N_{bits}} \cdot V_{sens}} \cdot R_{digital} = \frac{5.0\text{V}}{2^{24} \cdot 0.05\frac{\text{V}}{\text{Pa}}} \cdot R_{digital} \qquad (2)$$

So the equivalent pressure value in mPa, R_{mpa}, becomes a function of the input in fixed-point integer, $R_{digital}$; whereby, $R_{mPa} = 5.9605 \times 10^{-3} \times R_{digital}$.

Noise floor recordings were taken by measuring the RMS value of a zero input to the ADC, and using Eqn. 2 to convert to pressure. The noise floor of the probe and amplifier is nominally -15 dB SPL, and the noise floor of the Sharc EZ-Kit was measured to be -14 dB SPL, for a combined system noise floor of about -8.5 dB SPL, or about 7.5 μPa. This noise floor will be reduced even further through the process of synchronous averaging.

During subject testing, the investigator makes parametric adjustments to the stimulus type (such as pulse width, rate, number of epochs, etc.) in a MATLAB GUI. The GUI communicates to the Sharc EZ-Kit through via UART. A C-header file is updated with the stimulus information, and is loaded via serial port into the Sharc EZ-Kit real-time (see Fig. 3). The on-board program memory triggers the stimulus on the DAC and acquires the amplified recording on the ADC which passes through a hardware DC-blocking filter, built-in to the AD1835. The signal is synchronously averaged in a data memory buffer, which is passed back to MATLAB for post-processing.

Upon acquiring the averaged and scaled responses, the signals were digitally filtered forwards and backwards to prevent phase distortion. A 4-pole Butterworth filter with cutoff frequencies of 0.3 and 8 kHz was used.

C. Stimuli

The acoustic clicks were generated by 5-sample wide (104.2 µs) rectangular pulses. DNLR responses were recorded at 82 dB-SPL (252 mPa), 76 dB-SPL (126 mPa), and 70 dB-SPL (63 mPa). A mixed amplitude train of 4 clicks was presented to each subject. Since the intensities of each 4-click sequence vary, then it is imperative to reconcile the nominal intensity of the stimuli with the effective averaged intensity. The described intensity level used throughout the manuscript corresponds to the base-to-peak equivalent sound pressure level, or dB pe-SPL as defined in [16]. In order to calibrate the effective intensity of the 4 click average, the 3 rarefaction clicks are a factor of -3.5 dB lower than the nominal intensity, and the condensation click is a factor of 6 dB greater than the nominal intensity. The resultant nonlinear-mode average is equal in intensity to a linear-mode average of 2 clicks.

4096 sweeps of DNL sequences, equivalent to 2048 linear sweeps, were synchronously averaged using mean averaging. The averaging window was 25 ms, or 1200 samples at a 48 kHz sampling rate. The stimuli were presented to the subjects at a rate of 39.1 Hz in order to reduce the impact of electromagnetic interference through the process of averaging. Artifact rejection of 5 mPa (48 dB SPL) was used during the time window of 5 to 25 ms post-stimulus onset.

D. Analysis of artifact size

The process of DNLR artifact reduction is not possible without the utility of a high bit-depth TEOAE acquisition system. The saturating part of the MR in a medium bit-depth system can never be removed through subtraction or any other linear process. Suffice to say that the MR of a 24-bit system will necessarily be smaller than in a 16-bit system since the entire MR is susceptible to linear subtraction (see Figs. 2 and 4 for a visual representation).

In this study, the artifact size of a 24-bit DNLR acquisition was compared to that of a standard 16-bit.

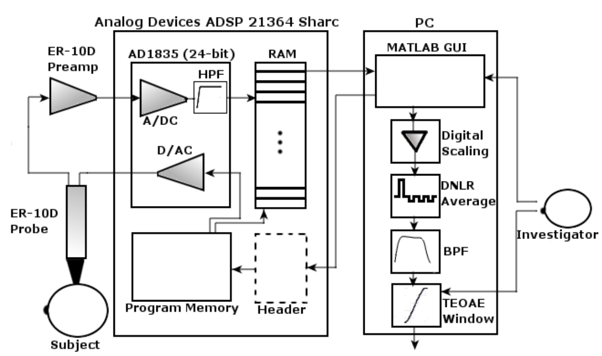

Fig. 3. High resolution OAE system block diagram. The investigator chooses the stimulus properties from a GUI, and this information is sent to the AD Sharc EZ-Kit via UART. The main program acquires from the ADC, using the DAC as a trigger. After a DC-Blocking filter, the response is synchronously averaged in a data memory buffer. After recording, the signal is passed back to MATLAB for post-processing, which includes scaling, DNLR averaging, zero phase-distortion BPF, and windowing.

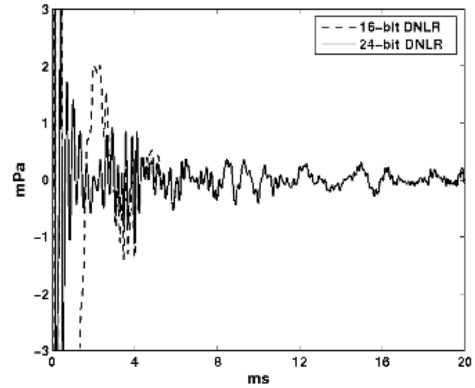

Fig. 4. DNLR response from a subject (04). A representation of a filtered, but unwindowed TEOAE. The solid line shows a 24-bit DNLR recovery, and the dashed line shows a 16-bit DNLR recovery. The 24-bit system has a greatly reduced artifact since the MR does not saturate the ADC.

Therefore, in order to study the effects of artifact reduction, an estimate of the artifact size must be obtained. This is determined by a peak-to-peak measurement, or the range, of the artifact. The range of the artifact-reduced MR can be compared to the range of the MR acquired using a standard DNLR technique to obtain a percent reduction in range, given by Eqn. 3.

$$\%\text{Reduction} = \frac{\text{range}[MR_{16-bit}] - \text{range}[MR_{24-bit}]}{\text{range}[MR_{16-bit}]} \cdot 100\% \quad (3)$$

E. Analysis of artifact duration

An objective method for analyzing the duration of the MRs was selected in order to compare 16-bit to 24-bit DNLR acquisition. The MPEG-7 standards delineate a specific temporal descriptor, which is intended to identify the effective duration of a transient [17]. This measure first derives the temporal envelope estimation by the process of moving a selectable window across the calculated energy of the signal, which is given by Eqn. 4.

$$e[n] = \sqrt{\frac{1}{N}\sum_{i=n-N+1}^{n} MR[i]^2} \quad (4)$$

The temporal envelope estimation is based on the short-time average energy, where the 1 ms window, N, acts as a low-pass filter of the energy, and n is the digitized time index. The effective duration, e, is the amount of time that the envelope is above a given threshold. A threshold of 1 mPa was chosen, so as not to exclude any TEOAEs, and to choose a level low enough such that the MR will not obstruct the TEOAE response significantly.

III. Conclusions

The results of this study show that a high bit-depth acquisition system can significantly improve the characteristics of the TEOAE. As shown in the results, the size of the meatal artifact acquired by the DNLR method is drastically reduced compared to medium or low bit-depth systems. The reduction of the stimulus artifact should, and does, lead to a reduction in the MR duration. This reduction helps reveal the early latency TEOAEs, which are expected to be high frequency (>4 kHz). It is thought that even a subtle improvement in MR duration can greatly improve the recovery of early latency TEOAEs. Exponential models, eg. those derived in [18] and [19], indicate that the frequency content of the early latency TEOAEs will increase greatly even with an analysis window that only marginally approaches the stimulus onset.

At stimulus intensities of 70, 76, and 82 dB SPL, the mean MR effective duration for the DNLR technique is below that of the tested DNLR technique, with improvements of 0.22 ms, 0.38 ms, and 0.53 ms respectively (see Figs. 5 and 6b). These values represent statistically significant improvements for intensity values of 76 and 82 dB-SPL.

The percent reduction of the artifact amplitude is shown in Fig. 6a. At relatively low stimulus levels (63 mPa), the reduction in stimulus artifact is modest (<23%). However, as the stimulus intensity increases, so does the reduction in the artifact, such that at 82 dB SPL (252 mPa) there is a 93% reduction in stimulus artifact size. The improvement in artifact amplitude range reduction is significant at all three stimulus intensities, with a T-test value of $p<0.01$

In conventional systems, the high-frequency (early latency) portion of the TEAOEs has been generally discarded due to the obstruction of the meatal artifact. There are several hearing disorders in which the recording of the high-frequency information provides diagnostic and therapeutic value. It is noted that after the treatment of patients with the drug Cisplatin, ototoxic effects are readily observable [20]. Ototoxic monitoring methods have been devised to allow for a moderated dosage in order to prevent sensorineural hearing loss. One clinical application for the recovery of higher TEOAE frequency content could be the monitoring of the ototoxic effects of Cisplatin. Further such studies may elucidate the effects of this drug on very high frequency cochlear functioning. The effectiveness of these methods currently has been hampered by the lack of information regarding high frequencies. The high resolution system described in this study provides those capabilities.

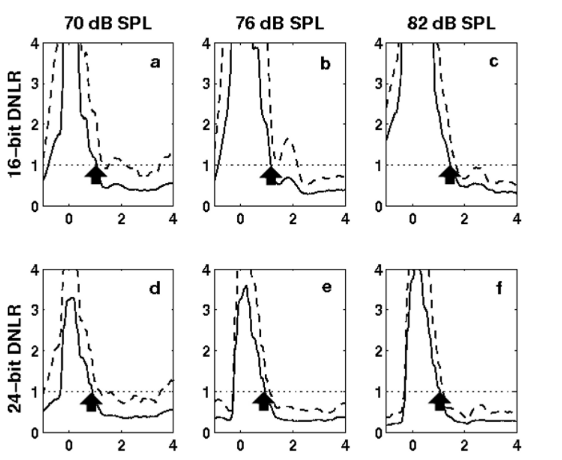

Fig. 5. Average MR effective duration. For 16-bit ADCs, the amplitude of the temporal envelope increases with stimulus intensity, causing an increase in effective duration at a 1 mPa threshold. However, with 24bit ADCs, the temporal envelope stays relatively constant, with a consistently lower average effective duration. Solid lines: mean, dashed lines: mean + std dev, and dotted lines: threshold. The arrow indicates the offset boundary of the average artifact.

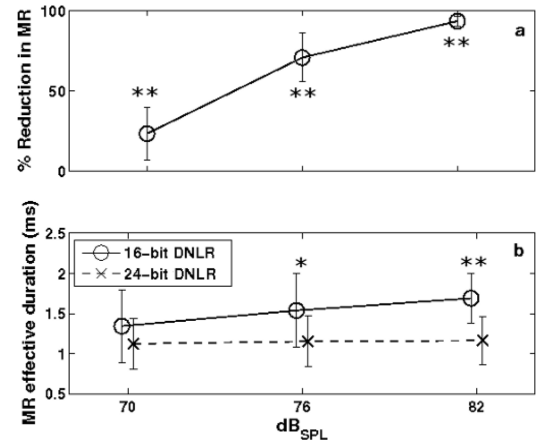

Fig. 6. (a) Percent reduction in MR using 24-bit mode. As the stimulus intensity increases, the improvement of the 24-bit mode increases as well, given as a percentage of the artifact amplitude. (b) MR duration for 16-bit and 24-bit acquisition modes. The 24-bit mode has a shorter MR duration than the typical 16-bit mode at all three tested stimulus levels. (Statistical significance: * indicates $p<0.05$ and ** indicates $p<0.01$).

IV. References

[1] D. Kemp, “Stimulated acoustic emissions from within the human auditory system,” *J Acoust Soc Am*, vol. 64, pp. 1386-91, 1978.

[2] P. Dallos, “The active cochlea,” *J Neurosci*, pp. 4575-85, 1992.

[3] C. Shera and J. Guinan Jr, “Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs,” *J Acoust Soc Am*, vol. 105, pp. 782-98, 1999.

[4] A. Maxon et al., “Using transient evoked otoacoustic emissions for neonatal hearing screening,” *Brit J Audiol*, vol. 27, pp. 149-53, 1993.

[5] P. M. Tujal and M. N. Souza, “Simulus Artifact Cancellation in Click Evoked Otoacoustic Emissions using Linear Prediction,” in *2003 IEEE Proc 20th Annu Conf EMBSand CMBEC*, pp. 3218-21.

[6] J. Liu, et al, “A new method for eliminating stimulus artifact in TEOAE using ICA,” *Lect Notes Comput Sc*, pp. 1041-47, 2004.

[7] M. Yuwaraj and H. Kunov, “Wavelet based reduction of stimulus artifact in evoked otoacoustic emissions testing,” in *1995 IEEE Proc 25th Annu Conf EMBS*, pp. 1055-56.

[8] R. Arslan, O. Ozdamar and Y. Ulgen, “Digital subtraction method for transient evoked otoacoustic emissions recording with ipsilateral noise suppression,” *Audiology*, vol. 40, pp. 55-62, 2001.

[9] D. Kemp, S. Ryan and P. Bray, “Otoacoustic emission analysis and interpretation for clinical purposes,” *Adv Audiol*, pp. 77-98, 1990.

[10] D. Kemp and P. Bray, “An advanced cochlear echo technique suitable for infant screening,” *Brit J Audiol*, vol. 21, pp. 191-204, 1987.

[11] D. Kemp, et al, “Acoustic emissions cochleography-practical aspects,” *Scand Audiol Supp*, vol. 25, pp. 71-95, 1986.

[12] D. Keefe, “Double-evoked otoacoustic emissions,” *J Acoust Soc Am*, vol. 103, pp. 3489-98, 1998.

[13] E. Zwicker, “A model describing nonlinearities in hearing by active processes with saturation at 40 dB,” *Biol Cybern*, pp. 243-250, 1979.

[14] B. Lineton, A. Thornton and V. Baker, “An investigation into the relationship between I/O nonlinearities and rate-induced nonlinearities of CEOAEs recorded using MLS,” *Hearing Res*, pp. 24-35, 2006.

[15] D. Kemp and R. Chum, “Properties of the generator of stimulated acoustic emissions,” *Hearing Res*, vol. 2, pp. 213-32, 1980.

[16] R. Burkard, “Sound pressure level measurement and spectral analysis of brief acoustic transients,” *Electroen Clin Neuro*, pp. 83-91, 1984.

[17] G. De Poli and L. Mion, “Algorithms for sound and music computing,” in *From audio to content*. Padova: U Padova, 2006, ch. 5.

[18] R. Sisto and A. Moleti, “On the frequency dependence of the otoacoustic emission latency in hypoacoustic and normal ears,” *J Acoust Soc Am*, vol. 111, pp. 297-308, 2002.

[19] G. Tognola, F. Grandori, and P. Ravazzani, “Time-frequency distributions of CEOAEs,” *Hearing Res*, vol. 106, pp. 112-22, 1997.

[20] D. McAlpine and B. Johnstone, “The ototoxic mechanism of cisplatin,” *Hearing Res*, vol. 47, pp. 191-203, 1990.