



## OPEN ACCESS

## EDITED BY

Jian Huang,  
Huazhong University of Science  
and Technology, China

## REVIEWED BY

Yao Wang,  
Tianjin Polytechnic University, China  
Yun Lu,  
Huizhou University, China

## \*CORRESPONDENCE

Shixiong Chen  
sx.chen@siat.ac.cn  
Yuan Tao  
taop19772000@163.com

†These authors have contributed  
equally to this work

## SPECIALTY SECTION

This article was submitted to  
Neural Technology,  
a section of the journal  
Frontiers in Neuroscience

RECEIVED 14 August 2022

ACCEPTED 26 September 2022

PUBLISHED 17 October 2022

## CITATION

Wang X, Zhu M, He Y, Liu Z, Huang X,  
Pan H, Wang M, Chen S, Tao Y and Li G  
(2022) Usefulness of phase gradients  
of otoacoustic emissions in auditory  
health screening: An exploration with  
swept tones.  
*Front. Neurosci.* 16:1018916.  
doi: 10.3389/fnins.2022.1018916

## COPYRIGHT

© 2022 Wang, Zhu, He, Liu, Huang,  
Pan, Wang, Chen, Tao and Li. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Usefulness of phase gradients of otoacoustic emissions in auditory health screening: An exploration with swept tones

Xin Wang<sup>1,2,3†</sup>, Mingxing Zhu<sup>4†</sup>, Yuchao He<sup>1,3</sup>, Zhenzhen Liu<sup>5</sup>,  
Xin Huang<sup>6</sup>, Hongguang Pan<sup>7</sup>, Mingjiang Wang<sup>4</sup>,  
Shixiong Chen<sup>1,3\*</sup>, Yuan Tao<sup>6\*</sup> and Guanglin Li<sup>1,3</sup>

<sup>1</sup>CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, <sup>2</sup>Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China, <sup>3</sup>Guangdong-Hong Kong-Macao Joint Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, <sup>4</sup>School of Electronics and Information Engineering, Harbin Institute of Technology, Shenzhen, China, <sup>5</sup>Surgery Division, Epilepsy Center, Shenzhen Children's Hospital, Shenzhen, China, <sup>6</sup>Department of Otorhinolaryngology, Peking University Shenzhen Hospital, Shenzhen, China, <sup>7</sup>Department of Otolaryngology, Shenzhen Children's Hospital, Shenzhen, China

Otoacoustic emissions (OAEs) are low-level sounds generated by the cochlea and widely used as a noninvasive tool to inspect cochlear impairments. However, only the amplitude information of OAE signals is used in current clinical tests, while the OAE phase containing important information about cochlear functions is commonly discarded, due to the insufficient frequency-resolution of existing OAE tests. In this study, swept tones with time-varying frequencies were used to measure stimulus frequency OAEs (SFOAEs) in human subjects, so that high-resolution phase spectra that are not available in existing OAE tests could be obtained and analyzed. The results showed that the phase of swept-tone SFOAEs demonstrated steep gradients as the frequency increased in human subjects with normal hearing. The steep phase gradients were sensitive to auditory functional abnormality caused by cochlear damage and stimulus artifacts introduced by system distortions. At low stimulus levels, the group delays derived from the phase gradients decreased from around 8.5 to 3 ms as the frequency increased from 1 to 10 kHz for subjects with normal hearing, and the pattern of group-delay versus frequency function showed significant difference for subjects with hearing loss. By using the swept-tone technology, the study suggests that the OAE phase gradients could provide highly sensitive information about the cochlear functions and therefore should be integrated into the conventional methods to improve the reliability of auditory health screening.

## KEYWORDS

swept-tone, otoacoustic emissions, cochlear impairment, phase gradients, auditory health screening

## Introduction

Otoacoustic emissions (OAEs) are low-level sounds generated by the normal activities of the cochlea and can be recorded by a sensitive microphone inside the ear canal. Although it is difficult to physically examine the cochlea due to its deep location inside the temporal bone, the discovery of OAEs provides a non-invasive window to observe the functional status of the cochlea. Studies have shown that OAEs are byproducts of the mechano-electrical activities of outer hair cells (OHCs) that can provide energy feedback to boost the vibrations of the basilar membrane and to amplify the cochlear response to incoming sounds (Dallos et al., 1997). Such physiological OHC activities, also called electromotility, are crucial for the extraordinary frequency selectivity and hearing sensitivity of the human auditory system (Liberman et al., 2002). During the active process of OHC electromotility, part of the extra energy provided by OHCs travels inversely along the basilar membrane and propagates to the ear canal and is recorded as OAEs. A study showed that OAE signals would significantly decrease or even disappear as a result of the blockage of the OHC electromotility (Brownell, 1990). Therefore, the presence of OAEs is a reliable indicator of thriving OHC activities as well as a normal functioning cochlea. Moreover, OAE signals are easy to measure and not affected by attentions or consciousness of the patients (Meric and Collet, 1994). Therefore, OAE measurements have been intensely used in routine hearing screening and audiological assessments in the clinic, especially for the pediatric population that is difficult to test in conventional audiogram assessments (Kemp et al., 1990).

Otoacoustic emission signals are commonly analyzed in the frequency domain by examining both amplitude and phase spectra. Since OHCs at different cochlear positions generate OAEs of different frequencies, their amplitude spectrum has been widely used as a direct approach to inspect the presence of OAE signals and to evaluate OHC functionalities at different frequencies. The OAE amplitude spectrum demonstrates regular spectral periodicity that is unique to each subject (Neumann et al., 1994; Talmadge et al., 1999; Wagner et al., 2008). Another distinctive feature of OAE signals is that the phase changes dramatically with frequency and there is a steep phase gradient in the OAE phase spectrum (Dhar et al., 2002; Choi et al., 2008; Martin et al., 2009; Henin et al., 2011). The steep phase gradient, which reflects the round-trip travel time of OAE signals, is closely related to the active process of OHC electromotility and the frequency selectivity of the cochlea (Shera and Bergevin, 2012). The unique feature of phase gradient enables it to act as an important tool for various studies in the peripheral auditory system (Shera and Guinan, 1999, 2003; Kalluri and Shera, 2001). The phase gradient could also be utilized to investigate the sources and generation mechanisms of different types of OAEs (Zweig and Shera, 1995;

Shera and Guinan, 1999; Kalluri and Shera, 2001; Goodman et al., 2003; Lineton and Lutman, 2003), to estimate the cochlear tuning by deducting group delays of the basilar membrane at different frequencies (Shera et al., 2002; Siegel et al., 2005; Shera and Bergevin, 2012), and to examine the olivocochlear efferent control of OHC activities introduced by a contralateral stimulus (Giraud et al., 1995; Guinan et al., 2003). Moreover, a recent study showed that there was a close relation between unstable OAE phase shift and a stiff cochlear partition, suggesting that the phase could be possibly used as a non-invasive way to detect endolymphatic hydrops of Menière's disease (Avan et al., 2011).

Although the OAE phase gradients are useful in different ways, they are mostly restricted to auditory research only and rarely used in clinical practices (Abdala et al., 2018; Liu et al., 2020). Currently, two types of OAEs are measured in the clinics: transient evoked otoacoustic emissions (TEOAEs) measured with brief tones such as clicks, and distortion product otoacoustic emissions (DPOAEs) induced by two sinusoids with closely spaced frequencies. However, only the OAE amplitude is provided upon the completion of both types of measurements while the phase information which essentially represents the OAE signals is completely discarded. One reason is that the measure of phase information requires that the OAEs be tested with a sufficient frequency resolution so that the phase difference between two neighboring frequencies does not exceed  $2\pi$  to avoid possible phase discontinuities (Shera et al., 2002). However, the frequency resolution of current OAE measurements is usually in the order of hundreds of hertz and it is far from sufficient to capture OAE phases that change dramatically with frequency (Shera and Bergevin, 2012). Recently, Chen et al. (Chen et al., 2013; Jun et al., 2014) proposed a method of using swept tones with rapidly varying frequencies to measure OAE signals with a frequency-resolution as high as a few hertz, making it a great candidate to measure the phase of OAE signals across frequencies. Another possible reason for the lack of use of OAE phases is that the phases of TEOAEs and DPOAEs could be deteriorated by their complex generation mechanisms (Shera and Guinan, 1999). The multiple reflections of TEOAEs (Kemp et al., 1990; Avan et al., 1993; Tognola et al., 1997) and two distinctive sources of DPOAEs (Shera and Guinan, 1999, 2003; Kalluri and Shera, 2001) make it rather difficult to interpret the phase information of the two types of OAEs currently used in the clinic. Stimulus frequency otoacoustic emission (SFOAE) is another type of OAEs commonly evoked by one single stimulus and it attracts increasing attention recently due to its appealing features when compared with other types of OAEs (Guinan et al., 2003; Choi et al., 2008; Bentsen et al., 2011; Cheatham et al., 2011; Ren et al., 2013). Studies showed that SFOAEs were more frequency-specific in reflecting the functional status of corresponding OHCs (Guinan et al., 2003; Choi et al., 2008). It is widely accepted that SFOAEs are generated by linear coherent reflections within the cochlea (Shera and Guinan, 1999;

Goodman et al., 2003; Lineton and Lutman, 2003), and therefore the interpretation of SFOAEs is less complicated, by avoiding multiple reflections in TEOAEs and source mixing in DPOAEs.

The purpose of this study is to use the swept-tone method to measure the phase gradients of SFOAEs in high frequency-resolution, and to demonstrate the role of using phase gradient in combination with conventional amplitude spectrum to obtain more reliable results of auditory health screening in the clinic.

## Materials and methods

### Subjects

Sixteen subjects were recruited from Shenzhen Institutes of Advanced Technology, with ages from 23 to 36 years old. All subjects declared that they had no congenital auditory disease in the family and no history of otological surgery or ototoxic drug usage. A conventional audiogram test was also performed on each subject at standard frequencies (1, 2, 3, 4, 6, and 8 kHz) prior to the experiment. Twelve subjects demonstrated normal hearing with thresholds better than 20 dB HL across all frequencies, and the other four subjects had mild hearing loss over certain frequencies due to long-term exposure to loud sounds. The subjects were told to lie comfortably on a foam-covered bed in a double-walled sound booth during the experimental tests. All subjects gave informed consent and provided permission of their data for scientific purposes. The protocol of this study was approved by the Institutional Review Board of Shenzhen Institutes of Advanced Technology (SIAT-IRB-130124-H0015).

### Equipment

The presentation of the stimuli and the recording of the acoustic response were controlled by a custom Windows PC program implemented in Matlab environment (Mathworks Inc., USA). Signal Processing Toolbox in Matlab was adopted in this manuscript to further analyze the signals. The digital waveforms of the stimuli were initially synthesized from the PC and then sent to an USB sound card (E-MU 0204, Creative Technology Ltd.) with very low noise background and rather low nonlinear distortion *via* a universal ASIO driver. The sound card is a full duplex with two input channels and four output channels, with all channels of pristine resolutions of 24 bits and very high signal-to-noise ratios (SNRs) up to 117 dB. Two ER-2A earphones (Etymotic Research) were connected to the sound card and converted the digital voltage to acoustic sounds to stimulate the auditory system at the ear canal. The acoustic response was simultaneously recorded by an ER-10+ low-noise microphone (Etymotic Research) seated together with the two earphones inside a foam earplug, and digitized by the USB

sound card at a sample rate of 48 kHz. All the original raw data were stored for offline analyses.

## Experimental design and procedures

### Three-interval protocol to extract stimulus frequency otoacoustic emissions

Since SFOAEs share the same frequency as the evoking stimulus, the extraction strategy becomes more complicated than the simple spectral analysis of TEOAEs or DPOAEs (Kalluri and Shera, 2007). A three-interval protocol (Keefe, 1998; Keefe and Ling, 1998; Chen et al., 2013; Jun et al., 2014) based on the two-tone suppression phenomenon (Kemp and Chum, 1980) was used to extract the SFOAEs in the experiment.

In the three-interval protocol, two stimulus tones ( $s_1$  and  $s_2$ ) were presented in a specially designed order within three intervals of equal duration (Figure 1A). When either stimulus ( $s_1$  or  $s_2$ ) was presented alone in the first or the second interval, the acoustic response ( $p_1$  or  $p_2$ ) contained both the stimulus artifact and the evoked SFOAEs. However, when the same stimuli  $s_1$  and  $s_2$  were presented simultaneously during the third interval, the SFOAEs evoked by either stimulus in the response  $p_{12}$  would be suppressed by the other stimulus and the amplitude would decrease by  $\Delta p_1$  or  $\Delta p_2$  as a result. In contrast, the stimulus artifacts related with  $s_1$  and  $s_2$  remained unchanged. When subtracting the response of  $p_{12}$  from  $p_1 p_2$ :

$$\Delta p = p_1 + p_2 - p_{12} \quad (1)$$

Most of the stimulus artifacts would be canceled out, and only the SFOAE amplitude changes in the third interval ( $\Delta p_1$  and  $\Delta p_2$ ), as well as other background noises and interferences, would be left in the residue  $\Delta p$ .

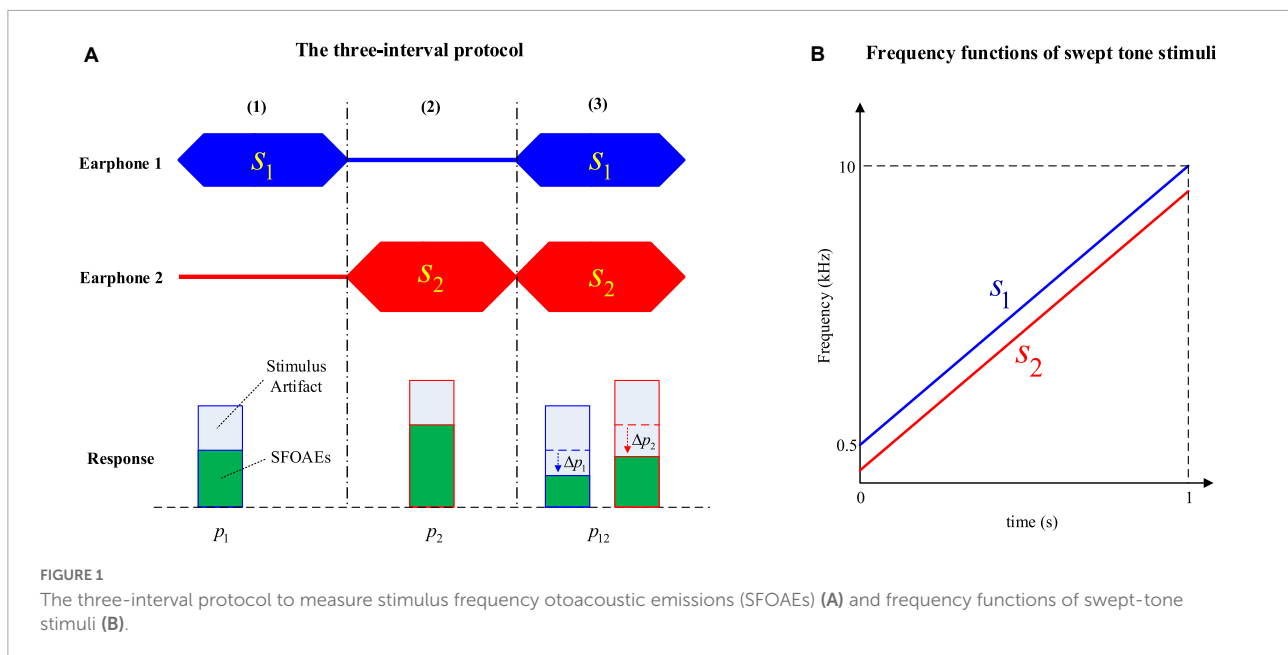
### Stimulus generation and presentation

In this study, both stimuli  $s_1$  and  $s_2$  were swept tones with time-varying frequencies to improve the efficiency of SFOAE measurements (Figure 1B). The swept tones were constructed by customizing the amplitude and phase spectra in the frequency domain, and converting to the time domain to get the temporal waveform *via* an inverse fast Fourier transform (iFFT) (Müller and Paulo, 2001; Chen et al., 2013).

In the experiment, the duration of all three intervals was kept at 1 s. The frequency of  $s_1$  (the probe tone) was increased linearly from 0.5 to 10 kHz within 1 s, and the frequency of  $s_2$  (the suppressor tone) was kept 200 Hz lower than  $s_1$  (Figure 1B). The level of  $s_1$  ( $L_1$ ) was increased from 45 to 60 dB FPL at a 5-dB step, and the level of  $s_2$  remained constant at 80 dB FPL.

### Experimental procedures

During the experiment, a foam earplug of the selected size was carefully inserted into the ear canal of the subject. Then



the digitally generated swept tones were presented by the two earphones in a three-interval fashion as shown in **Figure 1**, and the acoustic response at the ear canal was simultaneously collected by the microphone seated inside the earplug. For each signal condition, the same stimulus tones were repeatedly presented for 20 times and the responses were digitally averaged to improve the SNR. In the experiments, the stimulus level was equalized across frequencies according to the forward sound pressure (Chen et al., 2014) to avoid the impacts of standing waves. As a comparison, another session with no stimulus level calibration, during which there might be excessive stimulus level within a certain frequency range, was also carried out to examine the effects on the swept-tone SFOAEs.

For verification purposes, the earplug was also inserted into a plastic uniform tube with one end closed and the same procedures were performed to measure the response. The tube was 25 mm in length and 7 mm in diameter, approximately the same size as an adult ear canal. The acoustic response was recorded and analyzed to compare the differences between the human ear and the plastic tube.

## Data analysis

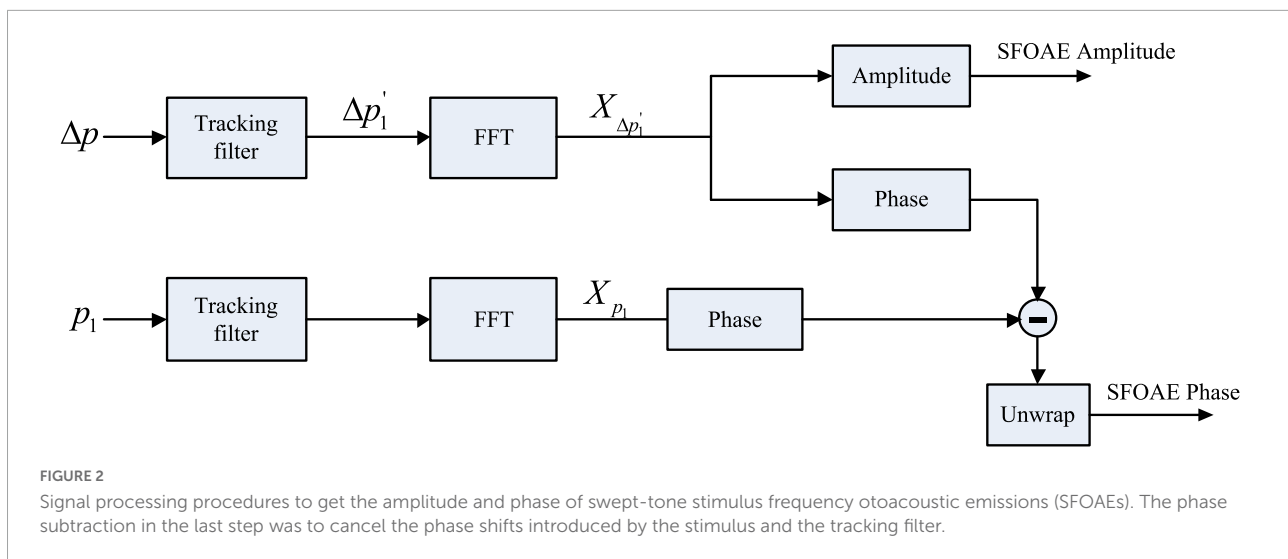
A tracking filter that can dynamically follow the instantaneous frequency of the target was used to extract the swept-tone SFOAEs in this study, and the details of the swept-tone and tracking filter could be found in Abdala et al. (2018) and Liu et al. (2020). Another advantage of the tracking filter is that it can easily attenuate the unwanted signal components by placing zeros around the corresponding frequency (Chen et al., 2013). As noted in Eq. 1, two major swept-tone SFOAE components ( $\Delta p_1$  and  $\Delta p_2$ ) coexisted in the residue response  $\Delta p$ . Since nonlinear system distortions

might be involved at high stimulus level (80 dB FPL) for  $s_2$ , only the swept-tone SFOAEs evoked by  $s_1$  (the  $\Delta p_1$  component) were analyzed in this study. Accordingly, there was one pole to track the  $\Delta p_1$  component, and one zero to attenuate the  $\Delta p_2$  component in the setup of the tracking filter. The signal passed through the tracking filter four times to improve the filter performance (Chen et al., 2013).

As shown in **Figure 2**, the tracking filter was applied to the temporal waveform of  $\Delta p$  to get a dynamic estimate ( $\Delta p_1$ ) of the swept-tone SFOAE component  $\Delta p_1$ . Then a fast Fourier transform (FFT), with a fixed 1-s length Hanning window, was performed on the  $\Delta p_1$ , and the magnitude of the FFT result  $X_{\Delta p_1}$  was taken as the amplitude spectrum of the swept-tone SFOAEs. Meanwhile, the center frequency of the tracking filter was set 100 Hz above the frequency of  $\Delta p_1$ , and the amplitude of the filter output was calculated as the reference of the noise floor.

To get the phase of the swept-tone SFOAEs, the same tracking filter that was used to obtain  $\Delta p_1$  was applied to the response of  $p_1$  in **Figure 1**, and an FFT was performed on the corresponding filter output to get the spectral complex  $X_{p_1}$  (**Figure 2**). Then the phase of  $X_{p_1}$  was subtracted from the phase of  $X_{\Delta p_1}$  (FFT result of  $\Delta p_1$ ), and the unwrapped phase difference was calculated as the phase of the swept-tone SFOAEs. The phase subtraction was used to eliminate the delays introduced by the recording system and the tracking filter, so that the phase gradient attribute to the swept-tone SFOAEs could be truly revealed.

Since the frequency of the stimulus continuously changed with time, the evoked swept-tone SFOAEs were also continuous in frequency, making it possible to obtain the phase spectrum of SFOAEs in high definition to avoid possible discontinuities. For



an SFOAE phase spectrum  $\phi(f)$  prepared by the procedures in **Figure 2**, the group delay  $\tau$ , defined as the transit time of a signal through a system, could be calculated by:

$$\tau = -\frac{1}{2\pi} \frac{d\phi}{df} \quad (2)$$

The group delay of SFOAEs was a rather useful measure of the travel time of the OAE signals inside the cochlea, and it could provide quite useful information about the functional status of the OHCs and the sharp tuning of the cochlea (Shera and Guinan, 2003; Siegel et al., 2005). In this study, the group delays were calculated at discrete frequencies ( $f_i$ ) from 1 to 10 kHz (at a 1-kHz step) for all the subjects. For each discrete frequency  $f_i$ , the phase-frequency function from  $f_i-100\text{Hz}$  to  $f_i+100\text{Hz}$  was fitted with a straight line and the slope was used to calculate the group delay according to Eq. 2.

## Results

### Swept-tone stimulus frequency otoacoustic emissions in subjects with normal hearing

The presence of OAE signals is a distinctive feature of the healthy human ear that makes it different from other passive systems such as acoustic tubes. The use of swept tones made it possible to observe the OAE features in such a high definition. For comparison purposes, swept-tone SFOAEs were measured in both human ears with normal hearing and a plastic tube of similar sizes under the same signal conditions ( $L_1 = 50$  dB FPL) in this study. A typical comparison of the spectrogram (energy distribution as a function of time and frequency), amplitude and phase spectra between the two

responses (represented by  $\Delta p$  in Eq. 1) was shown in **Figure 3**. The most important finding was that the responses in the human ear and plastic tube showed dramatically different patterns. For the response in the human ear, two ascending lines, which reflected the energy concentrations in the residue response, were clearly observed in the spectrogram (**Figure 3A**). The two lines corresponded to the two SFOAE components ( $\Delta p_1$  and  $\Delta p_2$ ) and had quite similar frequency patterns as their evoking stimuli (**Figure 1B**). Then a tracking filter was applied to extract the  $\Delta p_1$  component, and the filtered amplitude and phase spectra were shown in **Figures 3B,C**, respectively. The SFOAE amplitude in **Figure 3B** consisted of slow baseline variations and rapid spectral periodicity (or fine structures) indicated by alternating peaks and troughs. The overall SFOAE amplitude could get around 30 dB FPL above the noise floor for the subject with normal hearing. The pattern of the fine structures is unique for each specific subject. Another unique feature of the ear response was that the phase decreased dramatically as the frequency increased: the amount of the phase decrease could exceed 150 rad when the frequency increased from 0.5 to 10 kHz. However, for the response in the plastic tube, no swept-tone SFOAEs were observed in either the spectrogram (**Figure 3D**) or the amplitude spectrum (**Figure 3E**). The phase no longer demonstrated steep gradient and fluctuated around 0 as the frequency increased (**Figure 3F**).

### Swept-tone stimulus frequency otoacoustic emissions in subjects with mild hearing loss

Since SFOAEs are closely related to the normal functions of the cochlea, any alterations in the cochlear functions accompanied by auditory functional abnormality would result in changes in SFOAEs. A typical example of the swept-tone

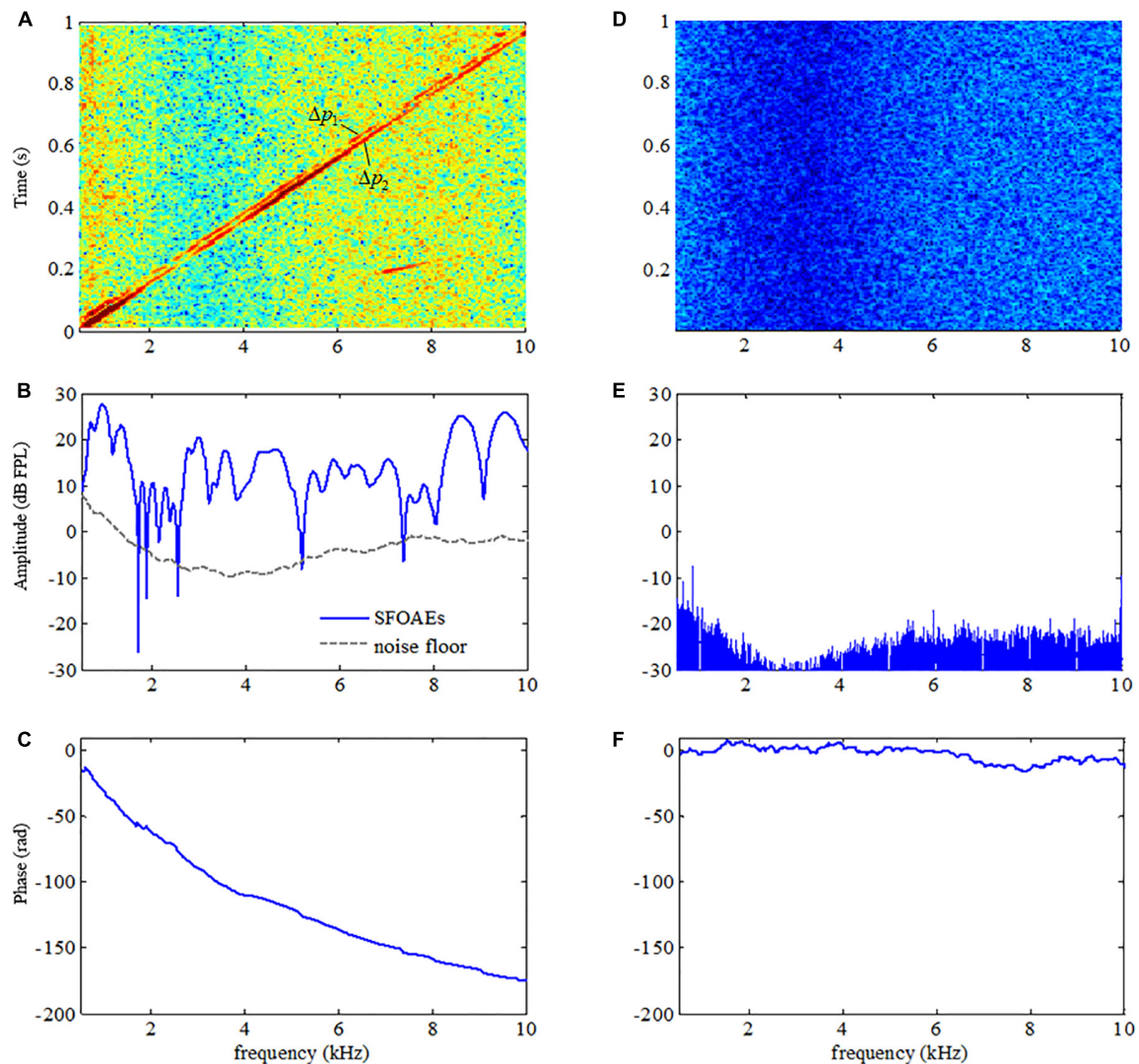


FIGURE 3  
Difference in the spectrograms, amplitude, and phase spectra between the acoustic responses in a human ear (A–C, respectively) and in a plastic tube (D–F, respectively).

SFOAEs of a subject with mild hearing loss of 2–3 kHz was shown in **Figure 4**. As shown in **Figure 4A**, the overall baseline amplitude of the SFOAEs fell below the noise floor within 2–3 kHz, which was consistent with the frequency region of the hearing loss. However, there was a large amplitude of SFOAEs over other frequencies. For the phase spectrum in **Figure 4B**, although steep phase gradients could be observed at most frequencies, the phase function became rather flat when the frequency was from 2 to 3 kHz. The flattening of the phase was consistent with the SFOAE amplitude reduction, as well as the region of the hearing loss. The abrupt phase discontinuities at other frequencies (such as 5–7.3 kHz) were due to the lower SNR at the trough of the SFOAE fine structures where the phase estimation was more susceptible to random noises.

## Phase gradients of swept-tone stimulus frequency otoacoustic emissions with artifacts

The SFOAEs of different frequencies originate from activities of OHCs at different positions along the cochlea. In practice, equalization of stimulus level across frequencies was usually desired so that SFOAEs from different frequencies could be compared. However, it is very common to have excessive stimulus level over certain frequency ranges without calibrations of stimulus levels, leading to large artifacts that could not be ignored in the SFOAE analyses (a typical example was shown in **Figure 5**). As shown in **Figure 5B**, a large amplitude of SFOAEs could be observed across all frequencies from 0.5

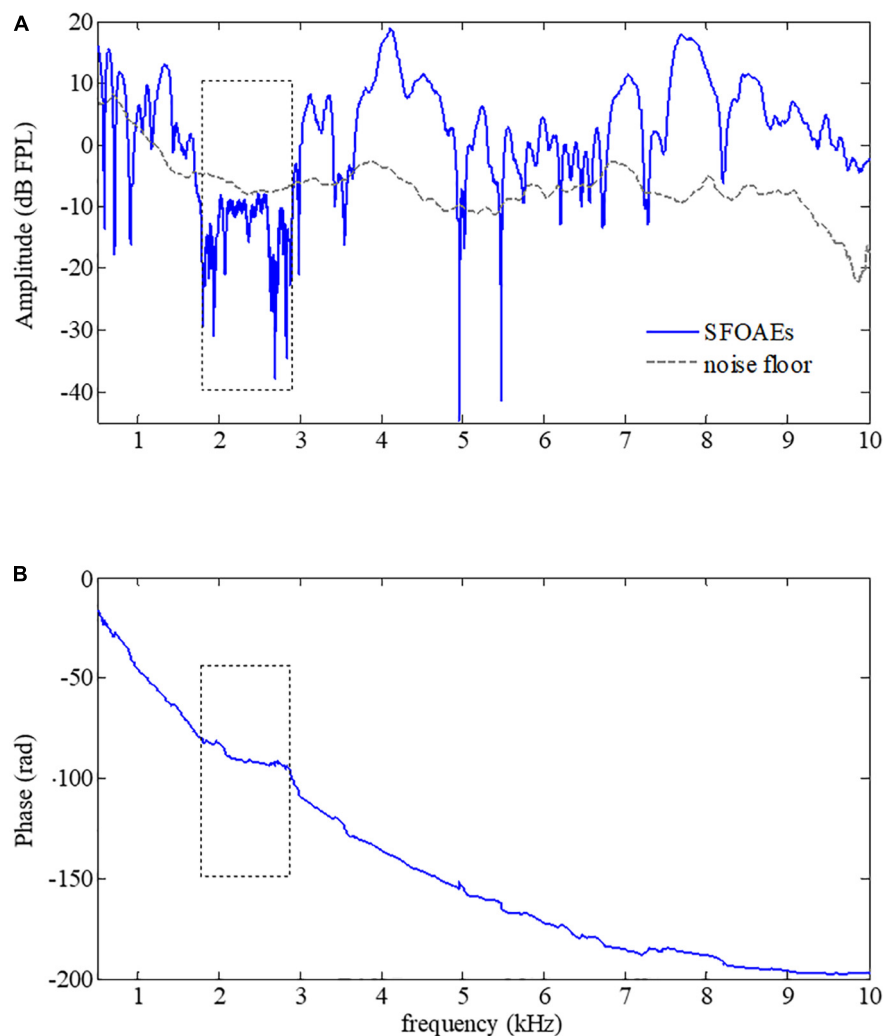


FIGURE 4

Amplitude (A) and phase (B) of swept-tone stimulus frequency otoacoustic emissions (SFOAEs) of a subject with hearing loss from 2 to 3 kHz (indicated by dotted boxes).

to 10 kHz, and it could reach up to 25 dB FPL around the frequency of 9 kHz. However, as we checked the corresponding stimulus level in [Figure 5A](#), it was found that the stimulus level above 5 kHz was much higher than the expected level of 50 dB FPL. Such excessive stimulus level (as high as 20 dB) could introduce nonlinear system distortions that could not be canceled out during the subtraction in Eq. 1 or the filtering by the tracking filter, resulting in unexpected artifacts during the SFOAE measurements ([Whitehead et al., 1995](#); [McCreery et al., 2009](#); [Burke et al., 2010](#); [Chen et al., 2014](#)). Fortunately, such artifacts could be correctly identified by the abnormal phase gradient in [Figure 5C](#), where the phase gradually violated the steep gradient pattern of OAE signals above 5 kHz. The flattening of the phase curve suggested that there might be a large amplitude of system distortion involved, leading to unreliable results in the extracted signals above 5 kHz.

## Groups delays of swept-tone stimulus frequency otoacoustic emissions

Group delays, obtained from the derivative of the phase versus frequency function (Eq. 2), could provide a non-invasive tool to monitor the cochlear tuning that is important for cochlear healthiness ([Shera and Bergevin, 2012](#)) as well as the frequency selectivity which is essential for speech perception ([Evans, 1975](#)). In this study, the SFOAEs were measured using swept tones to provide SFOAE phases in high frequency-resolution, making it possible to obtain reliable SFOAE group delays with high efficiency. In the experiment, the group delays of the swept-tone SFOAEs were measured at 10 discrete frequencies (from 1 to 10 kHz at a step of 1 kHz) under different stimulus levels. The means and standard deviations of the group delays averaged across the 12 normal-hearing

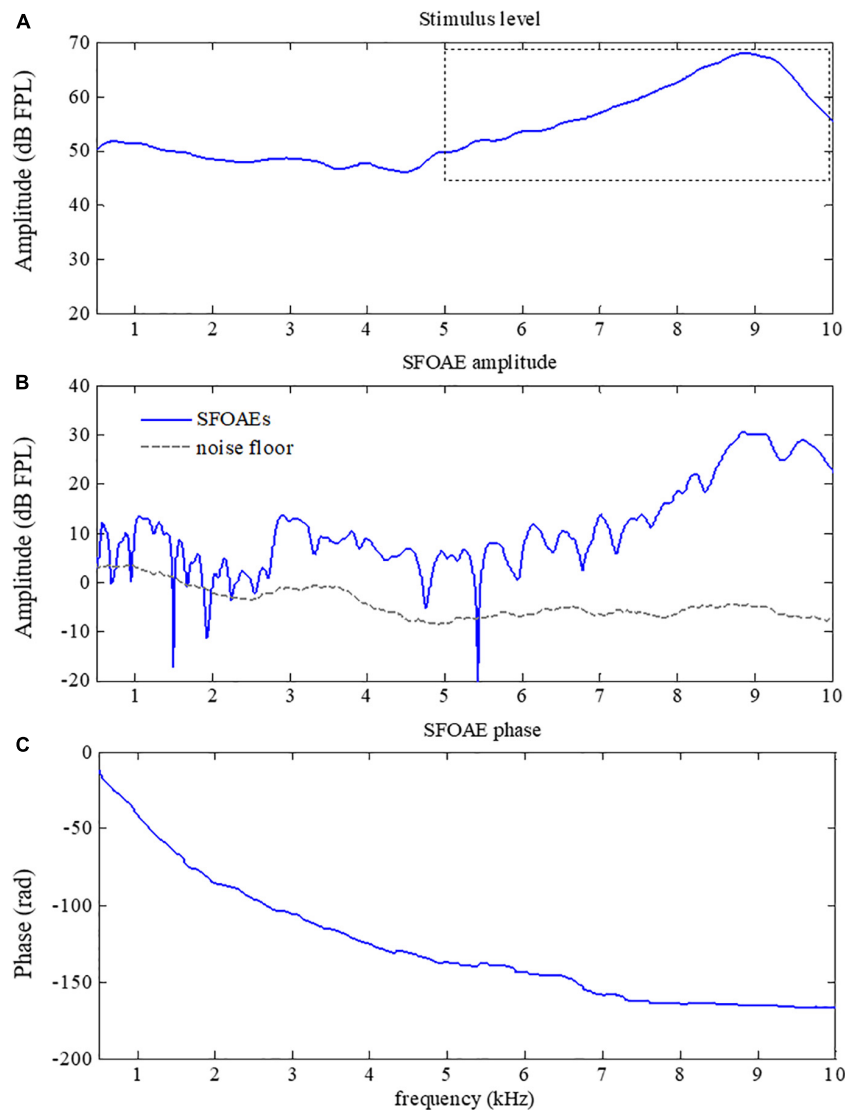


FIGURE 5

Amplitude (B) and phase (C) of swept-tone stimulus frequency otoacoustic emissions (SFOAEs) with artifacts above 5 kHz caused by excessive stimulus level (indicated by the dotted box in A).

subjects were shown in Figure 6. For the stimulus of 45 dB FPL, the average group delay decreased from around 8.5 ms to about 3 ms as the frequency increased from 1 to 10 kHz. The group delay decreased more rapidly at lower frequencies. When the stimulus level increased from 45 to 60 dB FPL, the group delay at a given frequency decreased monotonously, which was consistent with the findings of relevant studies (Shera and Guinan, 2003; Bentsen et al., 2011; Shera and Bergevin, 2012). Due to the impacts of low-frequency noises in the ear canal, the group delays of different stimulus levels almost overlapped at 1 kHz. Since the phase versus frequency function became rather flat as consequence of hearing loss (Figure 4), the estimated group delays within the corresponding frequency range would approach 0 ms and the group delay pattern would

be quite different from Figure 6 for subjects with hearing loss.

## Discussion

### Usefulness of stimulus frequency otoacoustic emission phase gradients in auditory health screening

The study showed that the phase gradients of swept-tone SFOAEs were rather useful to help improve the reliability of auditory health screening using OAE measurements. As



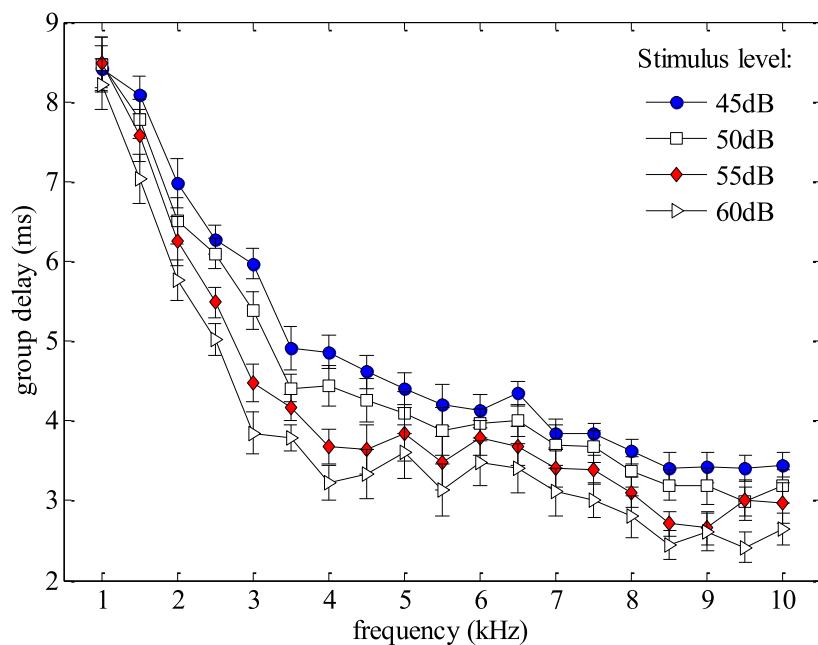


FIGURE 6

Group delays averaged across all subjects as functions of frequency and stimulus level derived from phase gradients of the swept-tone stimulus frequency otoacoustic emissions (SFOAEs).

observed in [Figure 3](#), the steep phase gradient of OAE signals is a unique feature of the normal functioning human cochlea. Such steep phase gradient would totally disappear if measurements were made in a passive tube ([Figure 3](#)) or partially vanish if auditory functional abnormality like hearing loss existed within certain frequency ranges ([Figure 4](#)). The close relation between the OAE phase gradients and cochlea healthiness is also reported in relevant studies using fixed-frequency tones ([Martin et al., 2009](#); [Avan et al., 2011](#); [Shera and Bergevin, 2012](#)). However, the phase information of OAEs is habitually abandoned due to insufficient frequency-resolution in clinical applications. Currently, only the results of OAE amplitude vs. noise level were provided and the presence of OAEs is determined by whether the OAE amplitude is above the noise level ([Kemp et al., 1990](#); [Erenberg et al., 1999](#); [Abujamra et al., 2013](#)). One problem is that the screening results are largely dependent on the reliability of the noise estimate. Since there is no universal standard for noise estimation, different levels of noise might be obtained if different algorithms are used for the noise calculation ([Attias et al., 2001](#); [Reavis et al., 2008](#)), leading to inconsistent results among different methods. In this study, the noise floor was obtained by setting the center frequency of the tracking filter 100 Hz above the swept-tone SFOAEs ([Figures 3–5](#)). However, slight differences in the estimated noise floor might be expected if the tracking filter was set in different ways, which might result in differences in the identification of a possible hearing-loss region in [Figure 4A](#). In contrast, the calculation of SFOAE phases is independent of the noise estimation, making it more

suitable and reliable to indicate the presence of OAE signals or the existence of possible auditory healthy issues (hearing loss in this study) over certain frequencies ([Figure 4B](#)). However, it is recommended that the phase of SFOAEs should be used in combination with the conventional amplitude spectra for more accurate auditory health screening in clinical practices. Moreover, the group delays calculated from the phase gradient could also be used to detect the abnormality of cochlear tuning that demonstrates evident alterations at the early stage of auditory functional disorders such as hearing loss ([Francis and Guinan, 2010](#); [Shera and Bergevin, 2012](#)).

## Phase gradients and group delays to detect otoacoustic emission artifacts

The present study also showed that the SFOAE phase gradients could help to identify possible artifacts that otherwise could be falsely treated as actual OAE signals. As mentioned earlier, the major difference between OAE signals and other responses is that the unique phase gradient ([Figure 3](#)) is so steep that the derived group delay is about 8.5 ms at 1 kHz for all subjects ([Figure 6](#)), not including the system round-trip delay that the stimulus spending on traveling along the outer and middle ears. The steep phase gradient, as well as the group delay, originates from the signal front delay which is the time difference between the onset of the basilar membrane (BM) and stapes, and the filter delay that the BM spends on

building the peak of the traveling wave (Ruggero, 2004). The filter delay of the BM, the major portion of the OAE group delay, is a unique physiological parameter closely related to the frequency selectivity of the cochlear tuning, with sharper tuning corresponding to longer group delays (Shera and Guinan, 2003; Ren et al., 2006). If there is no cochlear tuning involved (such as the response in a passive tube), the steep phase gradient would disappear (Figure 3) and the group delay would approach 0 as a consequence. In Figure 5, system distortions were involved above 5 kHz due to the excessive stimulus level, leading to incomplete cancellation of stimulus artifacts during the subtraction in Eq. 1. The remaining stimulus artifacts with zero group delay would dominate the low-level SFOAE components ( $\Delta p_1$  and  $\Delta p_2$ ) in the residual response  $\Delta p$  (Eq. 1), resulting in the violation of steep phase gradients at the corresponding frequency range (Figure 5C). However, it would be mistaken if we determined the presence of large OAE signals by merely checking the high SNR in the amplitude spectrum above 5 kHz (Figure 5B). Therefore, checking the steep phase gradient in combination with the amplitude spectrum is a more reliable way to distinguish OAE signals from other irrelevant interferences or unexpected noises.

## Conclusion

In this paper, SFOAEs were measured with swept tones in high frequency-resolution, so that the phase spectrum that is conventionally not feasible in auditory research or clinical tests could be obtained in a quite efficient way. The results demonstrated that the SFOAE phases in human ears showed steep gradients as the frequency increased, and such steep gradients are unique features that make the human-ear response different from other passive systems. The steep phase gradients could help to efficiently validate frequency regions of auditory functional abnormality, and to identify stimulus artifacts that could be mistakenly treated as evident OAE signals in practical applications. The pattern of the group delays derived from SFOAE phase gradients might be used to reflect the cochlear latency characteristics that were useful to evaluate the sharpness of the cochlear tuning and the normalcy of the cochlear frequency selectivity. The study suggested that using swept tones to measure SFOAEs and involving the phase information in combination with the amplitudes could be a rather promising approach to help improve the reliability of current hearing screening in the clinic.

## Data availability statement

The datasets presented in this article are available on reasonable request from the corresponding authors. Requests to access the datasets should be directed to SC.

## Ethics statement

The protocol of this study was approved by the Institutional Review Board of Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences (SIAT-IRB-130124-H0015). The patients/participants provided their written informed consent to participate in this study.

## Author contributions

SC and YT: conceptualization. XW and MZ: data curation. SC, GL, and MZ: funding acquisition. XW and SC: methodology, writing—original draft, and writing—review and editing. SC: supervision. XW, YH, and ZL: validation. XW, XH, and HP: visualization. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported in part by the Key-Area Research and Development Program of Guangdong Province (#2020B0909020004), National Natural Science Foundation of China (#81927804 and #62101538), Shenzhen Governmental Basic Research Grant (#JCYJ20180507182241622), Science and Technology Planning Project of Shenzhen (#JSGG20210713091808027 and #JSGG20211029095801002), China Postdoctoral Science Foundation (2022M710968), SIAT Innovation Program for Excellent Young Researchers (E1G027), and CAS President's International Fellowship Initiative Project (2022VEA0012).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Abdala, C., Ortmann, A. J., and Shera, C. A. (2018). Reflection- and distortion-source otoacoustic emissions: Evidence for increased irregularity in the human cochlea during aging. *J. Assoc. Res. Otolaryngol.* 19, 493–510. doi: 10.1007/s10162-018-0680-x
- Abujamra, A. L., Escosteguy, J. R., Dall'igna, C., Manica, D., Cigana, L. F., Coradini, P., et al. (2013). The use of high-frequency audiometry increases the diagnosis of asymptomatic hearing loss in pediatric patients treated with cisplatin-based chemotherapy. *Pediatr. Blood Cancer* 60, 474–478. doi: 10.1002/pbc.24236
- Attias, J., Horovitz, G., El-Hatib, N., and Nageris, B. (2001). Detection and clinical diagnosis of noise-induced hearing loss by otoacoustic emissions. *Noise Health* 3, 19–31.
- Avan, P., Bonfils, P., Loth, D., and Wit, H. P. (1993). Temporal patterns of transient-evoked otoacoustic emissions in normal and impaired cochleae. *Hear. Res.* 70, 109–120. doi: 10.1016/0378-5955(93)90055-6
- Avan, P., Giraudet, F., Chauveau, B., Gilain, L., and Mom, T. (2011). Unstable distortion-product otoacoustic emission phase in Meniere's disease. *Hear. Res.* 277, 88–95. doi: 10.1016/j.heares.2011.03.006
- Bentsen, T., Harte, J. M., and Dau, T. (2011). Human cochlear tuning estimates from stimulus-frequency otoacoustic emissions. *J. Acoust. Soc. Am.* 129, 3797–3807. doi: 10.1121/1.3575596
- Brownell, W. E. (1990). Outer hair cell electromotility and otoacoustic emissions. *Ear. Hear.* 11:82. doi: 10.1097/00003446-199004000-00003
- Burke, S. R., Rogers, A. R., Neely, S. T., Kopun, J. G., Tan, H., and Gorga, M. P. (2010). Influence of calibration method on distortion-product otoacoustic emission measurements: I. test performance. *Ear. Hear.* 31, 533–545. doi: 10.1097/AUD.0b013e3181d86b3d
- Cheatham, M. A., Katz, E. D., Charaziak, K., Dallos, P., and Siegel, J. H. (2011). Using stimulus frequency emissions to characterize cochlear function in mice. *AIP Conf. Proc.* 1403, 383–388. doi: 10.1063/1.3658115
- Chen, S., Deng, J., Bian, L., and Li, G. (2013). Stimulus frequency otoacoustic emissions evoked by swept tones. *Hear. Res.* 306, 104–114. doi: 10.1016/j.heares.2013.09.016
- Chen, S., Zhang, H., Wang, L., and Li, G. (2014). An in-situ calibration method and the effects on stimulus frequency otoacoustic emissions. *Biomed. Eng. Online* 13:95. doi: 10.1186/1475-925X-13-95
- Choi, Y. S., Lee, S. Y., Parham, K., Neely, S. T., and 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
- Dallos, P., He, D. Z. Z., Lin, X., Sziklai, I., Mehta, S., and Evans, B. N. (1997). Acetylcholine, outer hair cell electromotility, and the cochlear amplifier. *J. Neurosci.* 17, 2212–2226. doi: 10.1523/JNEUROSCI.17-06-02212.1997
- Dhar, S., Talmadge, C. L., Long, G. R., and Tubis, A. (2002). Multiple internal reflections in the cochlea and their effect on DPOAE fine structure. *J. Acoust. Soc. Am.* 112, 2882–2897. doi: 10.1121/1.1516757
- Erenberg, A., Lemons, J., Sia, C., Tunkel, D., Ziring, P., Adams, M., et al. (1999). Newborn and infant hearing loss: Detection and intervention. *Pediatrics* 103, 527–530. doi: 10.1542/peds.103.2.527
- Evans, E. (1975). The sharpening of cochlear frequency selectivity in the normal and abnormal cochlea. *Audiology* 14, 419–442. doi: 10.3109/00206097509071754
- Francis, N. A., and Guinan, J. J. (2010). Acoustic stimulation of human medial olivocochlear efferents reduces stimulus-frequency and click-evoked otoacoustic emission delays: Implications for cochlear filter bandwidths. *Hear. Res.* 267, 36–45. doi: 10.1016/j.heares.2010.04.009
- Giraud, A. L., Collet, L., Chery-Croze, S., Magnan, J., and Chays, A. (1995). Evidence of a medial olivocochlear involvement in contralateral suppression of otoacoustic emissions in humans. *Brain Res.* 705, 15–23. doi: 10.1016/0006-8993(95)01091-2
- Goodman, S. S., Withnell, R. H., and Shera, C. A. (2003). The origin of SFOAE microstructure in the guinea pig. *Hear. Res.* 183, 7–17. doi: 10.1016/s0378-5955(03)00193-x
- Guinan, J. J. Jr., Backus, B. C., Lilaonitkul, W., and Aharonson, V. (2003). Medial olivocochlear efferent reflex in humans: Otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs. *J. Assoc. Res. Otolaryngol.* 4, 521–540. doi: 10.1007/s10162-002-3037-3
- Henin, S., Thompson, S., Abdelrazeq, S., and Long, G. R. (2011). Changes in amplitude and phase of distortion-product otoacoustic emission fine-structure and separated components during efferent activation. *J. Acoust. Soc. Am.* 129, 2068–2079. doi: 10.1121/1.3543945
- Jun, D., Shixiong, C., Xiaoping, Z., and Guanglin, L. (2014). Using a dynamic tracking filter to extract distortion-product otoacoustic emissions evoked with swept-tone signals. *IEEE J. Biomed. Health Inf.* 18, 1186–1195. doi: 10.1109/JBHI.2013.2285558
- Kalluri, R., and Shera, C. A. (2001). Distortion-product source unmixing: A test of the two-mechanism model for DPOAE generation. *J. Acoust. Soc. Am.* 109, 622–637. doi: 10.1121/1.1334597
- Kalluri, R., and Shera, C. A. (2007). Comparing stimulus-frequency otoacoustic emissions measured by compression, suppression, and spectral smoothing. *J. Acoust. Soc. Am.* 122, 3562–3575. doi: 10.1121/1.2793604
- Keefe, D. H. (1998). Double-evoked otoacoustic emissions. I. Measurement theory and nonlinear coherence. *J. Acoust. Soc. Am.* 103, 3489–3498.
- Keefe, D. H., and Ling, R. (1998). Double-evoked otoacoustic emissions. II. Intermittent noise rejection, calibration and ear-canal measurements. *J. Acoust. Soc. Am.* 103, 3499–3508. doi: 10.1121/1.423058
- Kemp, D. T., and Chum, R. A. (1980). "Observations on the generator mechanism of stimulus frequency acoustic emissions – Two tone suppression," in *Psychophysical, physiological and behavioural studies in hearing*, eds G. van den Brink and F. A. Bilsen (Delft: Delft Univ. Press). doi: 10.1016/s0378-5955(98)00045-8
- Kemp, D. T., Ryan, S., and Bray, P. (1990). A guide to the effective use of otoacoustic emissions. *Ear. Hear.* 11, 93–105. doi: 10.1097/00003446-199004000-00004
- Lieberman, M. C., Gao, J., He, D. Z. Z., Wu, X., Jia, S., and Zuo, J. (2002). Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature* 419, 300–304. doi: 10.1038/nature01059
- Lineton, B., and Lutman, M. E. (2003). A parametric model of the spectral periodicity of stimulus frequency otoacoustic emissions. *J. Acoust. Soc. Am.* 114, 883–895. doi: 10.1121/1.1582176
- Liu, Y.-W., Kao, S.-L., Wu, H.-T., Liu, T.-C., Fang, T.-Y., and Wang, P.-C. (2020). Transient-evoked otoacoustic emission signals predicting outcomes of acute sensorineural hearing loss in patients with Meniere's disease. *Acta Otolaryngol.* 140, 230–235. doi: 10.1080/00016489.2019.1704865
- Martin, G. K., Stagner, B. B., Fahey, P. F., and Lonsbury-Martin, B. L. (2009). Steep and shallow phase gradient distortion product otoacoustic emissions arising basal to the primary tones. *J. Acoust. Soc. Am.* 125, EL85–EL92. doi: 10.1121/1.3073734
- McCreery, R. W., Pittman, A., Lewis, J., Neely, S. T., and Stelmachowicz, P. G. (2009). Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification. *J. Acoust. Soc. Am.* 126, 15–24. doi: 10.1121/1.3143142
- Meric, C., and Collet, L. (1994). Attention and otoacoustic emissions: A review. *Neurosci. Biobehav. Rev.* 18, 215–222. doi: 10.1016/0149-7634(94)90026-4
- Müller, S., and Paulo, M. (2001). Transfer-function measurement with sweeps. *J. Audio Eng. Soc.* 49, 443–471.
- Neumann, J., Uppenkamp, S., and Kollmeier, B. (1994). Chirp evoked otoacoustic emissions. *Hear. Res.* 79, 17–25. doi: 10.1016/0378-5955(94)90123-6
- Reavis, K. M., Phillips, D. S., Fausti, S. A., Gordon, J. S., Helt, W. J., Wilmington, D., et al. (2008). Factors affecting sensitivity of distortion-product otoacoustic emissions to ototoxic hearing loss. *Ear. Hear.* 29, 875–893. doi: 10.1097/AUD.0b013e318181ad99
- Ren, T., He, W., Scott, M., and Nuttall, A. L. (2006). Group delay of acoustic emissions in the ear. *J. Neurophysiol.* 96, 2785–2791. doi: 10.1152/jn.00374.2006
- Ren, T., Zheng, J., He, W., and Nuttall, A. L. (2013). Measurement of amplitude and delay of stimulus frequency otoacoustic emissions. *J. Otol.* 8, 57–62. doi: 10.1016/S1672-2930(13)50008-0
- Ruggero, M. A. (2004). Comparison of group delays of 2f1–f2 distortion product otoacoustic emissions and cochlear travel times. *Acoust. Res. Lett. Online* 5, 143–147. doi: 10.1121/1.1771711
- Shera, C. A., and Bergevin, C. (2012). Obtaining reliable phase-gradient delays from otoacoustic emission data. *J. Acoust. Soc. Am.* 132, 927–943. doi: 10.1121/1.4730916
- Shera, C. A., and Guinan, J. J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs. *J. Acoust. Soc. Am.* 105, 782–798. doi: 10.1121/1.426948

- Shera, C. A., and Guinan, J. J. (2003). Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning. *J. Acoust. Soc. Am.* 113, 2762–2772. doi: 10.1121/1.1557211
- Shera, C. A., Guinan, J. J., and Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc. Natl. Acad. Sci. U.S.A.* 99, 3318–3323. doi: 10.1073/pnas.032675099
- Siegel, J. H., Cerka, A. J., Recio-Spinoso, A., Temchin, A. N., van Dijk, P., and Ruggero, M. A. (2005). Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering. *J. Acoust. Soc. Am.* 118, 2434–2443. doi: 10.1121/1.2005867
- Talmadge, C. L., Long, G. R., Tubis, A., and Dhar, S. (1999). Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. *J. Acoust. Soc. Am.* 105, 275–292. doi: 10.1121/1.424584
- Tognola, G., Grandori, F., and Ravazzani, P. (1997). Time-frequency distributions of click-evoked otoacoustic emissions. *Hear. Res.* 106, 112–122. doi: 10.1016/S0378-5955(97)00007-5
- Wagner, W., Plinkert, P. K., Vonthein, R., and Plontke, S. K. (2008). Fine structure of distortion product otoacoustic emissions: Its dependence on age and hearing threshold and clinical implications. *Eur. Arch. Otorhinolaryngol.* 265, 1165–1172. doi: 10.1007/s00405-008-0593-0
- Whitehead, M. L., Stagner, B. B., Lonsbury-Martin, B. L., and Martin, G. K. (1995). Effects of ear-canal standing waves on measurements of distortion-product otoacoustic emissions. *J. Acoust. Soc. Am.* 98, 3200–3214. doi: 10.1121/1.413810
- Zweig, G., and Shera, C. A. (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. *J. Acoust. Soc. Am.* 98, 2018–2047. doi: 10.1121/1.413320