Using Swept Tones to Evoke Stimulus Frequency Otoacoustic Emissions

with In-situ Calibration

by

Shixiong Chen

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved May 2012 by the
Graduate Supervisory Committee:

Lin Bian, Chair
William Yost
Tamiko Azuma
Michael Dorman

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

Otoacoustic emissions (OAEs) are soft sounds generated by the inner ear and can be recorded within the ear canal. Since OAEs can reflect the functional status of the inner ear, OAE measurements have been widely used for hearing loss screening in the clinic. However, there are limitations in current clinical OAE measurements, such as the restricted frequency range, low efficiency and inaccurate calibration.

In this dissertation project, a new method of OAE measurement which used a swept tone to evoke the stimulus frequency OAEs (SFOAEs) was developed to overcome the limitations of current methods. In addition, an in-situ calibration was applied to equalize the spectral level of the swept-tone stimulus at the tympanic membrane (TM). With this method, SFOAEs could be recorded with high resolution over a wide frequency range within one or two minutes. Two experiments were conducted to verify the accuracy of the in-situ calibration and to test the performance of the swept-tone SFOAEs.

In experiment I, the calibration of the TM sound pressure was verified in both acoustic cavities and real ears by using a second probe microphone. In addition, the benefits of the in-situ calibration were investigated by measuring OAEs under different calibration conditions. Results showed that the TM pressure could be predicted correctly, and the in-situ calibration provided the most reliable results in OAE measurements.

In experiment II, a three-interval paradigm with a tracking-filter technique was used to record the swept-tone SFOAEs in 20 normal-hearing subjects. The

test-retest reliability of the swept-tone SFOAEs was examined using a repeated-measure design under various stimulus levels and durations. The accuracy of the swept-tone method was evaluated by comparisons with a standard method using discrete pure tones. Results showed that SFOAEs could be reliably and accurately measured with the swept-tone method. Comparing with the pure-tone approach, the swept-tone method showed significantly improved efficiency.

The swept-tone SFOAEs with in-situ calibration may be an alternative of current clinical OAE measurements for more detailed evaluation of inner ear function and accurate diagnosis.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

ix

Chapter 1

BACKGROUND AND SIGNIFICANCE

1.1 Project overview

The inner ear is a sensory organ that converts sound vibrations into neural signals. Within the inner ear, there are two types of hair cells: the inner hair cells (IHCs) and outer hair cells (OHCs). While the IHCs are mainly responsible for encoding sensory information, the OHCs provide feedback to the vibration of the inner ear structures to maintain an optimal stimulation of the IHCs. The key feature of OHC functions is that the mechanical feedback is compressive, with more energy fed back to the vibration for low levels of sound inputs. As byproducts of such OHC feedback, extra energies are generated from the inner ear. The generated energies propagate outwards and can be recorded as so-called otoscoustic emissions (OAEs) by a sensitive microphone placed inside the ear canal. Since OAEs are closely related to normal OHC activities, the measurement of OAEs is widely used in the clinic as a non-evasive tool for hearing loss screening.

Although OAE measurements have been used for over 30 years, the clinical applications of OAEs are still limited to hearing screening and no further information is available for hearing diagnosis. The purpose of this dissertation is to develop a new method to provide more detailed information about the functional status of the cochlea in an efficient way. The new method was to use a swept tone with the frequency sweeping continuously across a wide range and the energy equalized at the tympanic membrane (TM). This project was carried out in

1

two steps. The first step (chapter 2) was to test the performance and benefits of a calibration procedure to ensure the precise controlling of sound pressure at the TM. The second step (chapter 3) was to measure a type of OAEs using swept tones in human ears. The reliability of OAEs measured with the new method was investigated under various signal conditions, and the accuracy was cross-validated with the standard methods. The calibration procedure described in the first step was used to control the stimulus level during the OAE measurements in second step.

1.2 The auditory periphery

In order to understand the generation mechanisms and measurement principles of OAEs, it is necessary to briefly review the structures and functions of the auditory periphery.

The structures of the auditory periphery include three parts (Figure 1): the outer ear, the middle ear and the inner ear. The outer ear is composed of the pinna and the ear canal. The middle ear contains the tympanic membrane (TM) and the ossicles (the malleus, incus and stapes). The inner ear consists of mainly the cochlea that is the most important structure of the auditory periphery. Inside the cochlea, there is the basilar membrane which divides the fluid-filled space lengthwise into different canals, and the hair cells lie on top of the basilar membrane. The auditory nerves connect the hair cells with the structures in the central auditory system.
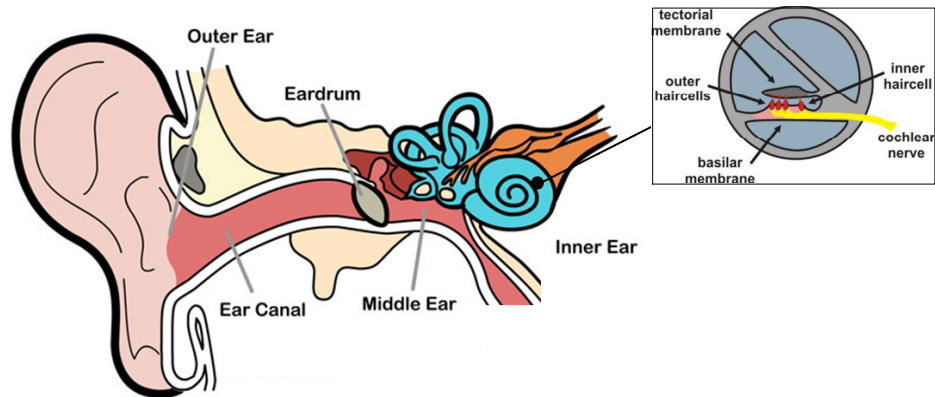
Figure 1. The main structures of the human auditory periphery (from sciencewithme.com).

When a sound comes, the acoustic energy is collected by the pinna and guided by the ear canal to strike the TM. The vibration of the TM then sets the ossicles into motion which in turn moves the fluid in the cochlea. The movement of the cochlear fluid causes the basilar membrane to vibrate. The hair cells on the basilar membrane convert the vibration into neural pulses that are carried by the auditory nerves to the brain to be interpreted as meaningful sounds.

1.2.1 Ear canal

It is crucial to understand the acoustics of the ear canal because OAEs are measured here. The ear canal can be considered as a tube with one end open to the air and the other end terminated by the TM. When the sound hits the TM, part of the sound energy is reflected back to the ear canal and the rest vibrates the TM and ossicles. The reflected sound wave can interact with the incident (original) wave and cause enhancements or cancellations, dependent on the phase relationship between the two sound waves. Such enhancements or cancellations form a standing wave in the ear canal, which can cause a non-uniform distribution

of the sound pressure along the length of the ear canal. At the entry of the ear canal, a pressure null is present when the incident and reflected waves are out of phase. Currently, most OAE equipment calibrates the stimulus level according to the sound pressure at the entry of the ear canal simply because it is where the probe microphone is located. However, such calibration can lead to inaccurate quantification of the stimulus level, given that the pressure at the TM can be significantly different (Siegel, 1994; Siegel & Hirohata, 1994). In the first experiment reported in Chapter 2, we attempted to address this issue by using an improved calibration method.

1.2.2 Middle ear

The middle ear also has significant effects on the OAE measurement since it is the pathway of the stimulus and OAEs. One primary function of the middle ear is to match the acoustic impedances between the air in the outer ear and the fluid in the inner ear, so that the sound energy can be effectively transmitted from the outer ear into the inner ear. The impedance match is achieved by the amplification of the pressure when a sound passes through the middle ear. The middle-ear amplification is frequency dependent with a resonant frequency around 1-4 kHz and it forms the shape of the human equal loudness contour (Ross, 1968). In this way, the middle ear can greatly affect the hearing sensitivity (Moore et al., 1997) and the spectral shape of OAEs (Suzuki & Takeshima, 2004). OAEs will be attenuated or even absent if the middle ear has problems, such as abnormal middle-ear pressures (Marshall et al., 1997) and middle-ear effusion (Nozza et al., 1997).

1.2.3 Basilar membrane

The basilar membrane (BM) of the inner ear is the place where OAEs originate. The BM is a spectral analyzer of incoming sounds. It goes along the length of the cochlea duct, and the morphological attributes differ from the base (close to the stapes) to the apex. The BM is narrow and stiff at the base, while it is wide and flexible at the apex. Therefore, each location has a characteristic frequency (CF), with high CFs located at the base and low CFs at the apex. When a sound enters the inner ear, the pressure difference within the cochlear fluid causes the BM to vibrate. The vibration initiates at the base and gradually builds up as the energy propagates towards the apex. The vibration reaches a peak at the place where the CF of the BM equals to the frequency of the incoming sound, and quickly dies out when it goes further. Such a sequence of vibration patterns of the BM resembles a wave moving along the cochlea and it is called the travelling wave. For a pure-tone stimulus, the peak of the travelling wave is located at the base for high-frequency input, and at the apex for low-frequency input.

1.2.4 Hair cells

Hair cells lie on the BM and are covered by the tectorial membrane. There are two types of hair cells in the cochlea: inner hair cells (IHCs) and outer hair cells (OHCs). There are one row of IHCs and three rows of OHCs. On top of both types of hair cells, there are hair-like projections called stereocilia with the tips embedded in the tectorial membrane. The travelling wave in response to a sound input causes the relative movement between the BM and tectorial membrane, which in turn causes the stereocilia to bend. The bending of the stereocilia opens

the ion channels to allow K^+ rush into the hair cells along the concentration

gradient. Such ion flow causes an electrical voltage change across the plasma

membrane. Thus the mechanical energy in sound is transformed into the receptor

potential of the hair cells (electrical energy), and this process is called the

mechano-electrical transduction (MET). One essential characteristics of the MET

process is its nonlinear transfer function, i.e., the relation between the mechanical

deflection of the stereocilia and the corresponding receptor potential is highly

nonlinear. The transduction is most sensitive at the resting point of the stereocilia

and it becomes compressive or saturated when the deflection increases in either

direction. The nonlinear characteristics of the transfer function can be described

by sigmoid-shaped functions, such as a hyperbolic-tangent function (Weiss &

Leong, 1985) or a Boltzman function (Patuzzi & Moleirinho, 1998).

Although both IHCs and OHCs are involved in the MET process, their

roles are different. The voltage change originated from the MET in IHCs leads to

the release of the neurotransmitters, and causes the afferent auditory nerve fibers

(ANFs) to fire. Actually, about 95% of the afferent ANFs contact with the IHCs,

making the IHCs the main sensory cells. In contrast, the voltage change from the

MET process in OHCs leads to the change of the cell length due to a voltage

dependent morphological change in the plasma membrane (Brownell et al., 1985).

In this way, OHCs can convert the voltage change (electrical energy) back to the

BM vibration (mechanical energy) and this process is called the reversed

transduction. The OHC mechanical response in a reversed transduction can

generate mechanical forces to help amplify the vibration of the BM to improve

6

the sensitivity of the IHCs. During the reversed transduction, some of the energy added by the OHCs is transmitted back from the inner ear to the outer ear, recorded as otoacoustic emissions in the ear canal.

1.3 Otoacoustic Emissions

Otoacoustic emissions (OAEs) were discovered by Kemp (1978) and became well-known phenomenon of the auditory system since then. OAEs can be recorded by placing a sensitive microphone in the ear canal, with or without presenting external stimuli.

The presence of the OAEs is strongly associated with the normal OHC activities (Hamernik et al., 1998; Glen K. Martin et al., 1998). If OHCs are damaged (such as by ototoxic drugs or by noise exposure), OAE amplitudes are decreased or even absent (Kossl & Vater, 2000; Subramaniam et al., 1994). Thus, OAEs provide a window to observe the functional status of the OHCs. Screening OAEs of different frequencies can help to identify the cochlear region where OHCs are damaged. Moreover, OAEs are relatively easy to measure, because it only requires an earphone to present the stimuli and a microphone to record the ear-canal response. In addition, measuring OAEs is an objective and yet non-invasive method of estimating the inner-ear function. Therefore, OAEs have been widely used as a means of audiological assessments in the clinic. OAE screening is an ideal way to screen the hearing loss in newborn babies who cannot respond behaviorally in standard hearing tests. The measurement of OAEs is useful in tracking ototoxicity during medical therapies using aminoglycoside antibiotics or chemotherapeutic drugs (Reavis et al., 2008; Stavroulaki et al., 2002) and

monitoring hearing functions during neuro-otologic surgeries (Telischi et al., 1995). OAEs are also used to assess the pathophysiological mechanisms of Meniere's diseases (van Huffelen et al., 1998). Combined with other hearing assessments, OAEs can assist clinicians in differential diagnosis to further pinpoint the cause of the hearing loss (Starr et al., 2001; Deltenre et al., 1999).

1.3.1 Types of OAEs

Depending on whether there is an evoking stimulus, all OAEs are divided into two distinct classes: spontaneous and evoked OAEs. Spontaneous OAEs (SOAEs) can be recordable in 80% of normal ears without any external stimuli (Penner & Zhang, 1997; Burns et al., 1992). Evoked OAEs can be further categorized into several different types based on the type of stimuli. If elicited by a transient stimulus (such as a click or a tone burst), the emissions are called transient evoked OAEs (TEOAEs); if two pure tones are used, the acoustic distortion products consisting of the combinations of the two evoking stimuli are recorded as distortion product OAEs (DPOAEs); if elicited by one single tone, they are referred to as the stimulus-frequency OAEs (SFOAEs).

Among the different types of evoked OAEs, TEOAEs and DPOAEs are currently used in the clinic. For TEOAEs, the transient can stimulate the entire length of the BM simultaneously and the emissions are observable after the stimulus. High-frequency TEOAEs are emitted by OHCs at the cochlear base with shorter latencies, followed by low-frequency TEOAEs that are generated from more apical regions with longer delays. TEOAEs are commonly used in hearing

8

screening because they are present as long as the OHCs are functioning, and they can be obtained efficiently.

DPOAEs are measured using two pure tones with closely spaced frequencies ($f_1, f_2; f_2 / f_1 \approx 1.2$). In normal ears, the recorded emissions contain a family of distortion products (DPs) in various combinations of integer multiples of the primary frequencies ($mf_1 \pm nf_2$). Among these components, the cubic difference tone ($2f_1 - f_2$) is used in the clinic because it is the most prominent DPOAE in humans (Telischi et al., 1995; Lonsbury-Martin et al., 1990). DPOAEs are widely used in auditory diagnosis because they are free from contamination by stimulus artifacts and they are faithful indicators of the nonlinearities of cochlear transductions (Bian, 2004; Bian et al., 2002).

1.3.2 Generation mechanisms

Even though OAEs are closely related to OHC functions, how they are generated differs depending on the stimulus and the signal level. Currently, there are two distinct generation mechanisms for evoked OAEs: linear coherent reflection and nonlinear distortion (Shera & Guinan, 1999). Linear coherent reflection views the OAEs being the energy reflected from impedance irregularities (or discontinuities) distributed randomly on the BM (Zweig & Shera, 1995). Several factors could be responsible for the impedance irregularities: the spacing of irregular or extra OHCs (Lonsbury-Martin et al., 1988), variations in the gains resulted from the OHC activities (Strube, 1989) and the change of the efferent innervations. The waves reflected by the irregularities at the peak region of the travelling wave have a coherent phase and the summed energy propagates

as a backward travelling wave. The OAEs generated by the linear reflection are characterized by steep phase gradients (Schairer et al., 2006). In contrast, the nonlinear distortion regards the OAEs as the energy originated from the nonlinear interactions of the travelling waves on the BM and it is highly dependent on the nonlinearities of the OHC activities. The nonlinear distortion is characterized by shallower phase gradients.

Different evoked emissions may involve different generation mechanisms. TEOAEs are wideband signals and they are thought to originate from linear reflections off preexisting impedance irregularities (Zweig & Shera, 1995). However, there are arguments that TEOAEs may involve nonlinear interactions among different frequency components of the emissions (Withnell & Yates, 1998). DPOAEs are generally considered as a mixture of two distinct components (Parazzini, Wilson, et al., 2006; Shaffer et al., 2003; Kalluri & Shera, 2001; Shera & Guinan, 1999). One component is the distortion from the overlapped region of the two stimulus-tone travelling waves and the other is from the reflection at the CF place ($2f_1 - f_2$ for the cubic difference tone). SFOAEs are considered to be generated from linear reflection at low to moderate stimulus levels, and nonlinear distortions may also be involved at high stimulus levels (Schairer et al., 2006; Goodman et al., 2003; Zweig & Shera, 1995).

1.3.3 Limitations of TEOAEs and DPOAEs

Although TEOAEs and DPOAEs are measured in the clinic for nearly 30 years, there are still limitations that impede their further applications in audiology. One limitation of TEOAEs is that there is a lack of high-frequency OAEs

10

particularly above 4 kHz. This is because the first 2-3 ms of the temporal

waveform of the acoustic response, which contains the short-latency high-

frequency OAEs, is greatly contaminated by the stimulus artifacts and therefore

removed from the analysis in TEOAE systems (Kemp et al., 1990). However,

studies showed that there are compelling reasons to measure high-frequency

emissions. In most cases, cochlear damage, such as ototoxicity (Schweitzer et al.,

1984; Brummett, 1980), noise induced hearing loss (Kuronen et al., 2003) and

other cochlear insults (Ohlemiller & Siegel, 1992), appear first at high frequencies

and then proceed to low frequencies. Therefore, monitoring of high-frequency

OAEs is useful for early detections of cochlear damage. Another limitation of

TEOAEs is that nonlinear interactions between different frequency components in

their generation mechanism may deteriorate the spectrum (Plinkert et al., 1999;

Wagner & Plinkert, 1999).

One limitation of DPOAEs is their complicated generation mechanisms.

They are considered as a mixture of two components from two distinct regions (as

mentioned previously), making it difficult to attribute the abnormality of

DPOAEs to the dysfunctionality of one specific generation region (Dhar et al.,

2002; Heitmann et al., 1996). Another limitation is that it tests only one frequency

at a time, making it less efficient than TEOAEs. Meenderink and van der Heijden

(2011) proposed to replace one of pure-tone stimuli with six to ten tones to evoke

DPOAEs. However, the choice of the stimulus frequencies is restricted by the

requirement of ensuring no DPs to be generated at the same frequency and

presenting multiple stimuli to a single earphone increases the chance of

overloading the transducer to produce distortion artifacts. Long et al (2008) used a continuous frequency-sweeping paradigm to evoke DPOAEs. This is an interesting approach since it can improve the efficiency of DPOAE measurement.

1.3.4 Advantages of SFOAEs

SFOAEs are evoked by one single tone and the emissions are generated at the same frequency as the stimulus. Unlike TEOAEs and DPOAEs involving complicated generation mechanisms from multiple regions, SFOAEs are thought to be generated by a simple linear reflection mechanism from a small cochlear portion located at the peak region of the traveling wave of the stimulus (Zweig & Shera, 1995). Thus, SFOAEs are considered to be very frequency-specific (Schairer et al., 2006; Goodman et al., 2003) and the abnormity of SFOAEs can map out the specific region of the damaged OHCs. The spectral distribution of SFOAEs across a wide frequency range could be a graphical representation of surviving OHCs which was coined by Kemp (1986) as the cochleogram. SFOAEs are vulnerable to cochlear damage induced by noise overexposure (Bentsen et al., 2011), and the group delay calculated from the phase spectrum can reflect functional status of OHCs (Shera & Guinan, 2003). Therefore, SFOAEs have advantages over other types of OAEs and show great potential to be developed as a clinical tool to be applied in the clinic.

However, SFOAEs are currently restricted to laboratory research only. One factor to prevent SFOAEs from clinical usage is that the emissions share the same frequency as their eliciting stimulus, making the extraction more difficult than other types of OAEs. Another reason is the low efficiency of the

measurement since SFOAEs are measured frequency-by-frequency using pure tones. As a result, only a narrow frequency range can be measured within the limited measurement time. It would be helpful if the frequency of the stimulus changes continuously with time, so that a wide range of frequencies can be tested within one single measurement.

1.4 Statement of purpose

The purpose of this dissertation is to develop a new method of OAE measurement to overcome the frequency limit of TEOAEs and to improve the efficiency of DPOAEs. The proposed method uses swept-tones whose frequency changes continuously with time to evoke SFOAEs that are more advantageous than other types of OAEs. By measuring SFOAEs evoked by swept tones, the method has the following advantages: 1) it is precise in specifying the location of the cochlear damage; 2) it is capable of measuring OAEs of any given frequency range; 3) the frequency resolution is very high since the swept-tone stimulus has a continuously frequency and a long duration; 4) it is highly efficient since it can test SFOAEs over a wide range of frequencies within one single measurement. Meanwhile, an in-situ calibration is also implemented to precisely control the level of the swept-tone stimulus at the TM position to resolve the standing wave issue that causes inaccurate stimulus calibration.

The hypotheses of this research are: 1) the in-situ calibration can precisely control the swept tone level at the TM and in turn improve the reliability of OAE measurements; 2) SFOAEs can be reliably measured with swept tones and the results are consistent with the SFOAEs measured with traditional methods using

pure tones. Two experiments are designed to test the two hypotheses, respectively. In the first experiment, the sound pressure at the TM was predicted by a multiple-cavity method and the predicted results were validated by measuring the actual sound pressure at the TM position using a second probe microphone. Then the predicted TM pressure was used in the in-situ calibration to equalize the acoustic energy across frequencies to evoke SOAEs, and the effects of the calibration on the OAE measurements were investigated. In the second experiment, the techniques of extracting swept-tone SFOAEs were implemented in human ears. The reliability of the swept-tone SFOAEs was investigated using a repeated measure design and the accuracy was examined by comparing the swept-tone SFOAEs with those obtained by a conventional pure tone methods.

Chapter 2

EXPERIMENT I

The calibration of the stimulus level refers to the adjustment of sound
energy that enters the middle-ear to be at the desired level. When the stimulus is a
wide-band signal, the calibration ensures that a desired spectrum is achieved. The
calibration of the stimulus level in OAE measurements is important because it
significantly affects the OAE amplitude and an inaccurate calibration can lead to
an unreliable result. The calibration in the real ear prior to each OAE
measurement is necessary, because the sound energy can vary depending on the
insertion depth of the probe, the seal of the ear tip, and the shape of the individual
ear canal. More importantly, for certain frequencies the sound pressure measured
at the entry of the ear canal can be drastically different from the pressure at the
TM. This is due to the presence of a standing wave inside the ear canal which is
formed by the enhancement and cancellation between the incident sound wave
from the ear phone and the reflected wave from the TM. A typical consequence of
the standing wave is that sound pressure measured at the ear canal entry shows a
deep notch at a certain frequency, but the sound level at the TM can reach a
maximal value. The standing wave problem occurs to the frequency that is
determined by the distance between the earphone probe and the TM. Therefore,
the reading from the probe microphone at the canal entry cannot reflect the actual
sound pressure at the TM, at least around the frequency of the standing wave.

For OAE measurement, a commonly accepted reference of the sound
energy entering the middle ear is the TM pressure which is measured with a

microphone placed within a few millimeters from the TM (Siegel, 1994; Siegel & Hirohata, 1994). However, inserting a microphone so close to the TM could cause discomfort and introduce potential threat of damaging the TM. As a result, measuring the TM pressure directly is rarely used in practice. The most common method is to calibrate the stimulus level according to the sound pressure measured from the probe microphone placed at the entry of the ear canal (called probe pressure). In such a calibration (namely the probe calibration), the driving voltage to the earphone is adjusted until the measured probe pressure achieves the desired spectrum. However, the probe calibration is largely affected by the standing waves in the ear canal. A notch is present at a certain frequency for the probe pressure but not for the TM pressure. If the probe pressure is equated across frequencies, the energy around the notch frequency would be over compensated. The over compensation can be as much as 15 to 20 dB around the notch frequency (Siegel & Hirohata, 1994). Whitehead et al (1995) used a "no calibration" strategy, in which the driving voltage to the earphone is held constant across frequency. However, the actual spectral shape at the TM is highly dependent on the frequency response of the earphone and the acoustic characteristics of the ear canal.

One solution for the calibration problem is to model the ear canal as a uniform tube (Dirks & Kincaid, 1987). While such modeling was successful in estimating the TM pressure below 6 kHz in some ear canals, several factors, such as the curvature of the ear canal and the inclination of the TM, led to increased estimation error of high frequencies. Other investigators tried to estimate the TM

16

pressure by measuring sound pressure at the TM position in the ear-canal

simulators of different lengths (Gilman & Dirks, 1986). They estimated the

distance between the probe and the TM by finding the spectral notch in the real-

ear probe pressure, and then placed the probe at the same position in the simulator.

The corresponding pressure measured at the TM position in the simulator was

used as the reference of the actual TM pressure. This may be a feasible calibration

method in practice. However, there is no guarantee that the acoustic

characteristics of the simulator are exactly the same as the real ear, given that

large individual differences exist in the ear-canal acoustics.

Recently, a calibration method based on estimating the incident pressure

(also called forward pressure) has been developed and is under intense

investigation (Burke et al., 2010; Rogers et al., 2010; Lewis et al., 2009; Withnell

et al., 2009; Scheperle et al., 2008; Neely & Gorga, 1998). This method tries to

separate the sound pressure at the entry of the ear canal into two components: the

incident and reflected pressures. It then adjusts the stimulus level according to one

individual component, the incident pressure, so that the standing-wave problem

caused by the interactions of the two components is avoided. However, the

separated pressure components cannot be validated since the incident or reflected

pressure alone cannot be actually measured anywhere in the ear canal.

The effects of various calibration methods on the OAE measurements

have been investigated in many studies (Burke et al., 2010; Rogers et al., 2010;

Scheperle et al., 2008; Whitehead, Stagner, et al., 1995; Siegel & Hirohata, 1994).

However, while some studies reported significant differences in OAE amplitudes

between different calibrations (Scheperle et al., 2008; Siegel & Hirohata, 1994),

little or no calibration effects were found in other studies (Burke et al., 2010;

Whitehead, Stagner, et al., 1995). One possible reason for such conflicting results

may be that the differences in the stimulus level between different calibration

methods may be present within a narrow frequency range, which may not be

covered by the sparse pure-tone frequencies measured by these studies. One

solution is to measure continuous-frequency OAEs so that any effects of the

calibration on the OAE amplitude can be observed in any frequency range.

In this experiment, an in-situ calibration was used to estimate and adjust

the sound pressure at the TM to a target level. The in-situ calibration is extended

from the incident-pressure method to attempt to predict the TM pressure. The in-

situ calibration can automatically adjust the driving voltage (as a function of

frequency) according to the predicted TM pressure until it reaches a desired

spectrum level. The predicted TM pressure was validated in a cylindrical acoustic

cavity and real ear canals by the actual measurement from a second probe

microphone. Furthermore, the benefits of the in-situ calibration in OAE

measurement were investigated by measuring the swept-tone SFOAEs (described

in the next chapter) with three different calibration methods that adjust the

stimulus level according to different sound pressures.

2.1 Methods

2.1.1 Materials and subjects

A) Materials

Two plastic tubes with a uniform diameter were used for different

purposes: a calculation tube and an evaluation tube. The calculation tube was

mainly used for estimating the acoustic properties of the earphone source, namely

the source pressure and source impedance. These parameters were used to predict

sound pressure at the TM. The inner diameter of the tube was 7 mm, which was

similar to the diameter of the human ear canal. The total length of the calculation

tube was about 80 mm, with one end open to be connected to the eartip and the

other end terminated by a piston. The piston could be moved to various positions

to set the tube at different effective lengths. The evaluation tube of 50 mm in

length was mainly used to validate the estimated impedance and predicted TM

pressure. For the evaluation tube, one end was open (to be connected with the

eartip) and the other end was closed by a flat plastic plate.
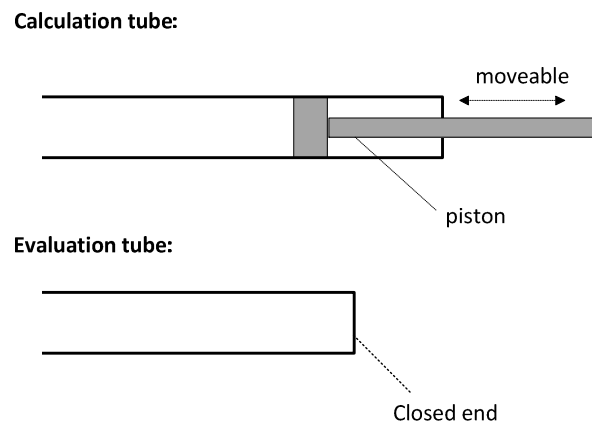
Figure 2. The materials used in this experiment: a calculation tube and an

evaluation tube.

B) Subjects

To validate the predicted TM pressure in real human ears and to examine

the effects of different calibrations on OAE measurement, five subjects (2 males,

3 females) with ages ranging from 21 to 30 years old (mean age=25) were

recruited in this experiment. The subjects had normal hearing with thresholds of

20 dB HL or less for frequencies from 250 Hz to 8k Hz. An otoscopic

examination was performed prior to the test to ensure a clean ear canal and a

normal functioning middle ear. The subjects were seated in a sound-proofed booth

comfortably and the experimental procedures were clearly explained to them. The

subjects were told to be as quiet as possible during the test. The recruiting

protocols were approved by the Institutional Review Board of Arizona State

University (ASU).

2.1.2 Equipment

The diagram of the sound delivering and recording system was shown in

Figure 3. A custom program developed in Labview (National Instruments, Austin,

TX) was used to generate the stimulus from a personal computer. The generated

digital signal was converted to analog voltage by a 24-bit signal acquisition and

generation card (PXI-4461, National Instruments). The voltage from the PXI-

4461 card was then used to drive an ER-2A earphone probe (Etymotic Research,

Elk Grove Village, IL) to present the acoustic stimulus. The earphone was seated

together with a low-noise ER-10B+ microphone (probe microphone) inside an

appropriately selected rubber eartip (ER-10D, Etymotic Research), which was

inserted into the open end of the tube or the ear canals. The sound pressure near

the earphone probe (probe pressure) was recorded by the probe microphone and then digitized by the PXI-4461 card at a sample rate of 80k sample/s.
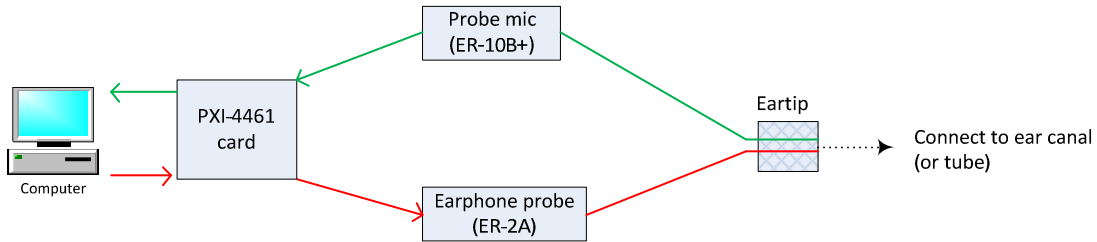
Figure 3. The diagram of the sound delivering and recording system.

2.1.3 Stimulus

A swept tone constructed in the frequency domain with a desired flat spectrum was used as the stimulus in this experiment (Appendix A). The swept tone was a frequency modulation tone whose frequency increased linearly with time. There were three signal parameters to control the swept tone stimulus: the frequency range (from f_1 to f_2), the duration (T) and the amplitude. A swept tone with a flat amplitude spectrum was initially generated. However, the recorded tube or ear canal frequency response was not flat due to the resonance and the standing wave in the cavities. The amplitude of the swept-tone was defined as the spectral level of the measured pressured at 1 kHz. For signal calibrations, the spectral levels of the swept tone for all frequencies were adjusted based on the inverse of the tube or canal frequency response to equalize the energy with respect to 1 kHz. The frequency response used in the calibration could be either the probe pressure or the predicted TM pressure, thus resulting in two type of calibration: probe and in-situ calibrations. In this experiment, the

signal parameters were: $f_1 = 0.5$ kHz, $f_2 = 10$ kHz, $T = 1$ s, amplitude = 50 dB SPL.

2.1.4 Procedures

A) Source calibration

The source parameters of the ER-2A earphone included the source pressure P_s and source impedance Z_s. The source pressure refers to the output pressure from the earphone without acoustic loads, and the source impedance is the obstruction of the earphone structures to the output pressure. The source calibration method used in this experiment was similar to Allen's multiple-cavity approach (1986), in which P_s and Z_s were solved by a transmission line model. The technical details of the multiple-cavity approach can be found in Appendix B and the procedures are described as followed.

The eartip with the earphone and probe microphone was tightly connected to the open end of the calculation tube and the piston was then set to five different positions. This resulted in 5 different tube lengths that ranged from 10 to 60 mm with roughly 10 mm apart. At each piston position, a swept-tone stimulus with a flat spectrum of 50 dB SPL was presented and the pressure response was recorded. The theoretical acoustic impedance of each tube length was calculated based on equation (8) in Appendix B. Then the sound pressure responses and the tube impedance functions of all the five piston positions were used to solve for P_s and Z_s in the frequency domain (see Appendix B).

22

To examine whether the choice of piston positions affected the solution, two sets of 5 piston positions were used to result in two sets of solutions of P_s and Z_s for a comparison.

B) Prediction of TM pressure

After P_s and Z_s were obtained in the source calibration, the eartip was moved to the evaluation tube or an ear canal to estimate the impedance and the TM pressure. Then the same swept tone stimulus was presented and the probe pressure at the eartip (P_L) was recorded. The impedance of the evaluation tube or the ear canal could be obtained in the frequency domain by:

$$Z_L = \frac{Z_s P_L}{P_s - P_L} \qquad (1)$$

With the impedance of the evaluation tube or the ear canal known, the incident pressure (P_+) and reflected pressure (P_-) at the entry of the ear canal could be obtained by (Rogers et al., 2010; Withnell et al., 2009; Scheperle et al., 2008):

$$P_+ = \frac{1}{2} P_L (1 + \frac{Z_0}{Z_L}) \qquad (2)$$

$$P_- = \frac{1}{2} P_L (1 - \frac{Z_0}{Z_L}) \qquad (3)$$

where $Z_0 = \rho c$ (ρ: air density; c: sound speed). Assuming that the absorption of the sound energy was negligible when the sound travelled from one end of the ear canal to the other end, the two components at the TM position [the incident pressure (P_+') and reflected pressure (P_-')] had the same amplitude as the two

23

components at the entry of the ear canal [P_+ and P_-] respectively but with different phases. Therefore, the P_+' and P_-' could be obtained by shifting the phase of P_+ and P_- (Appendix C). Finally, the predicted TM pressure ($\hat{P}_{TM}$) could be obtained by the sum of the two components at the TM position:

$$\hat{P}_{TM} = P_+' + P_-' \tag{4}$$

C) Verification of predicted TM pressure

It was essential to ensure that the predicted TM pressure ($\hat{P}_{TM}$) matched the actual TM pressure before it could be used for the calibration. The verification was conducted first in the evaluation tube then in the ear canals of the five human subjects (Figure 4). To verify the predicted TM sound pressure, a second probe microphone (Philips SHM3300) was routed down the tube or canal with a soft probe tube (ER7-14) so that the sound pressure at the closed end of the tube or at the TM could be recorded. The inner diameter of the probe tube was 1 mm and the length was 70 mm.

The insertion depth of the probe tube was determined by estimating the effective tube length or the ear canal depth from measuring the spectrum of the probe pressure. The entrance of the tube or ear canal was sealed with the eartip and the probe tube was adjusted to ensure the tip was with 1~2 mm from the closed end or the TM. The predicted TM pressure $\hat{P}_{TM}$ and the actual TM pressure measured by the second probe microphone (P_{TM}) were compared.
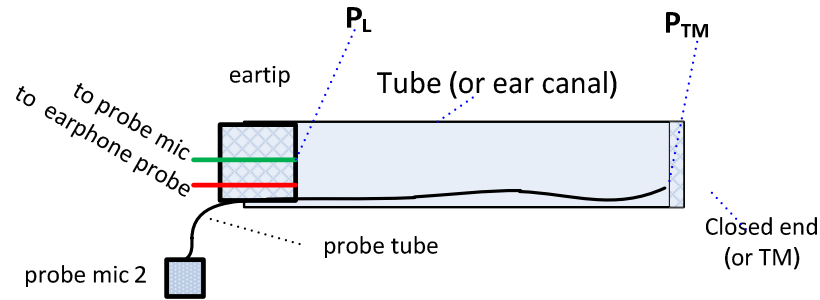
Figure 4. The verification of the predicted TM pressure in the tube and in the ear canal.

In order to compare $\hat{P}_{TM}$ and P_{TM}, the frequency responses of the two probe microphones were measured separately and the difference was used to correct the outputs from them (see Appendix D for details).

D) Effects of calibrations on OAE measurements

The calibration for OAE measurements refers to the process of adjusting the swept tone amplitude until the recorded sound pressure reaches a target which is a flat amplitude spectrum set at a desired level, i.e., 50 dB SPL. In this experiment, the stimulus was a swept tone with the duration of 1s and the frequency range from 0.5 to 10 kHz. The adjustment was done programmatically with a feedback loop. The acceptable range of the variation was set to less than 1 dB across frequency.

Depending on which signal was used for the adjustment, there were three different calibration conditions: no calibration, probe calibration and in-situ calibration. The duration of the swept tone stimulus was 1s, and the frequency range was from 0.5 to 10 kHz. For the "no calibration" condition, the amplitude (constant across frequency) of the swept tone was adjusted until the measured

25

probe pressure at 1 kHz reached 50 dB SPL. For the "probe calibration" condition, the swept tone was adjusted until the probe pressure matched a flat target of 50 dB SPL across frequency. For the "in-situ calibration" condition, adjustment was made to the swept tone until the predicted TM pressure reached a flat target of 50 dB SPL. For each subject, the swept-tone SFOAEs (described in the next chapter) were measured under the three calibrations successively. The SFOAE amplitudes were compared among the different calibration conditions.

2.2 Results

2.2.1 Source calibration

In order to predict the TM pressure, the acoustic properties (P_s and Z_s) of the earphone were determined by the source calibration. During the source calibration, the piston of the calculation tube was set to five different positions and the corresponding probe pressures were measured to solve for P_s and Z_s. To examine whether the choice of piston positions had effects on the solution, two sets of piston positions were used, resulting in two solutions of P_s and Z_s (Figure 5). The effective tube lengths of the first set of piston positions were 52.1, 43.2, 39.2, 16.6 and 8.9 mm, while the lengths of the second set were 60.2, 47.5, 35.3, 22.7 and 12.0 mm. It was observed that the two solutions of P_s and Z_s were in good agreement, indicated by the overlapped amplitude and phase spectra across the entire frequency range (0.5-10 kHz). The maximal deviation of P_s amplitude was 0.51 dB, and it was 0.72 dB for Z_s. The phase had a maximal difference of 0.06 rad for P_s and 0.14 rad for Z_s. The difference seemed to be larger at higher

26

frequencies. The comparison of the two solutions suggested that the selection of tube lengths did not affect the estimate of the source parameters P_s and Z_s.
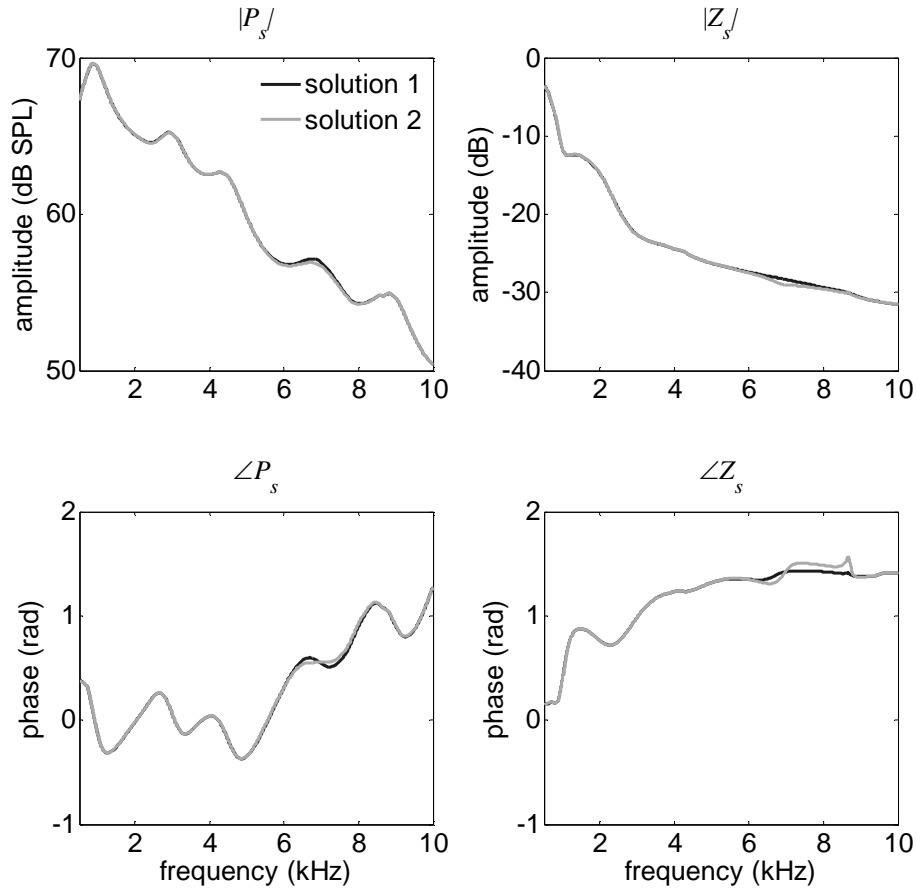
Figure 5. The amplitude and phase of the source parameters solved by two different sets of tube lengths.

2.2.2 Prediction in tube

Once source parameters (P_s and Z_s) were reliably obtained, they were then used to estimate the impedance of the evaluation tube or the real ear. The measurement in the evaluation tube provided a means to evaluate the impedance estimation, because the impedance of a tube can be calculated from theory

(Appendix B). The amplitude ($\left|Z_{tube}\right|$) and phase ($\angle Z_{tube}$) of the estimated and

theoretical impedance of the evaluation tube with an effective length of 33.6 mm

are shown in Figure 6. It was observed that the theoretical $\left|Z_{tube}\right|$ showed deep

notches at the frequency of $f_0 = 2.53\,\text{kHz}$ and its 3rd harmonics ($3f_0$). Meanwhile,

a peak could be found at the frequency of $2f_0$. For the phase $\angle Z_{tube}$, it changed

from $-\pi/2$ to $\pi/2$ at the notch frequencies, and from $\pi/2$ to $-\pi/2$ at the peak

frequencies. It was also observed that the estimated impedance from the

measurement matched the theoretical impedance well for both the amplitude and

the phase. The only noticeable difference was that the notch of the calculated

amplitude was not as deep, and the calculated phase changes at the notch (or peak)

frequency were not as abrupt. Nevertheless, the theoretical and estimated

impedances were in good agreement.
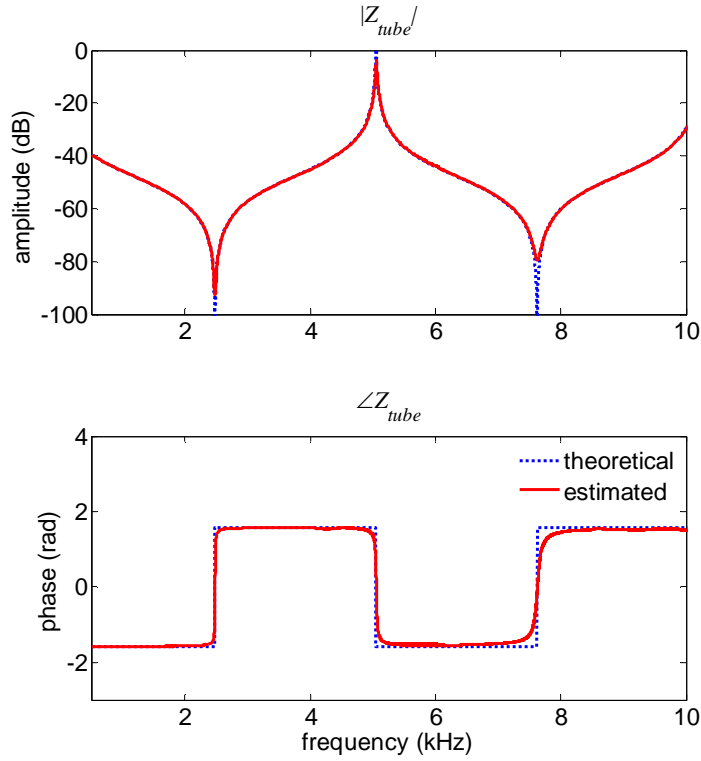
Figure 6. The amplitudes and phases of the estimated and theoretical impedances of the evaluation tube.

Based on the estimated impedance, the pressure at the closed end of the evaluation tube, equivalent to the TM position in the ear canal (Figure 4), could be predicted as $\hat{P}_{TM}$. Meanwhile, the actual pressure at the closed end (P_{TM}) was measured by a second microphone with a probe tube. The predicted $\hat{P}_{TM}$ and the actual P_{TM} are compared along with the probe pressure at the entrance (P_L) were compared in Figure 7. It could be observed that the probe pressure P_L had notches at about 2.53 kHz (f_0) and its third harmonic ($3f_0$). The notches could be as deep as 50 dB. Between the two notches, a pressure peak was observed at the

frequency around $2f_0$. These notches and peaks corresponded to the

measurement of $\left| Z_{tube} \right|$ in the same tube (Figure 6). In contrast, the deep notches

were not present in both the predicted and actual sound pressures at the closed end

of the tube (P_{TM} and $\hat{P}_{TM}$). In other words, the sound pressures of the two ends of

the tube differed substantially around the notch frequency and P_L could

underestimate the actual P_{TM} by as much as 50 dB in the tube. However, all of

these measurements showed a peak at about $2f_0$.

More importantly, the results showed that the predicted pressure $\hat{P}_{TM}$ and

the actual pressure P_{TM} were in good agreement across the entire frequency range

(0.5-10 kHz). The maximal prediction error, defined as $P_{TM} - \hat{P}_{TM}$, was less than

1.2 dB across frequencies. The prediction seemed to be best around the peak

frequency, indicated by the overlapped lines. The prediction error seemed to

slightly increase around the notch frequency and higher frequencies above 8 kHz.

In general, the results indicated that the prediction was very successful in the tube.
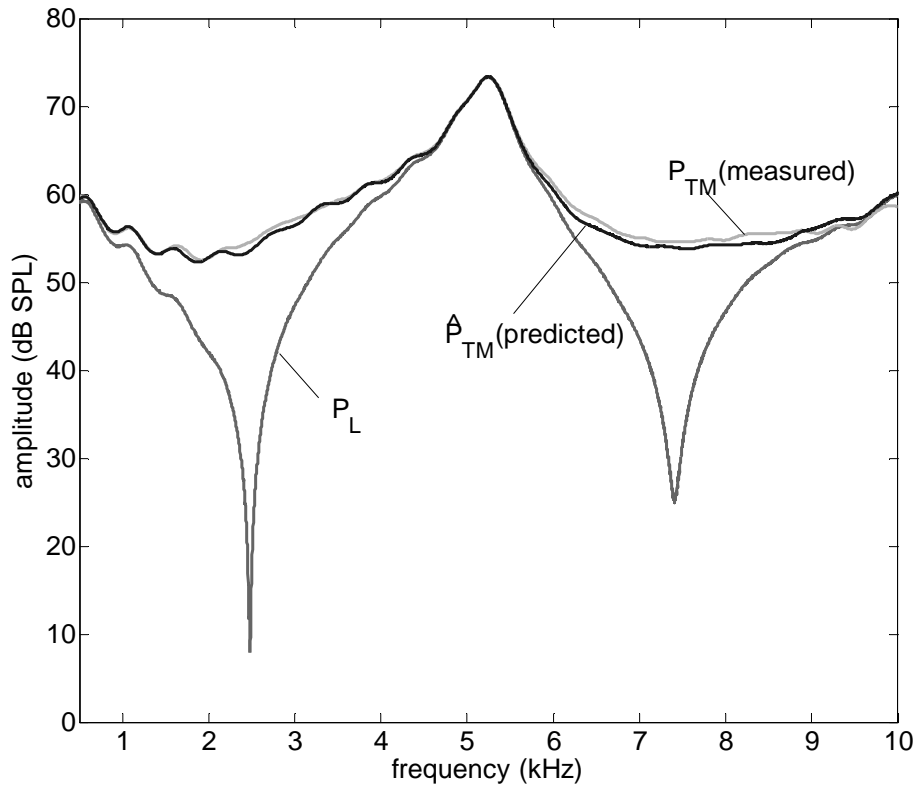
Figure 7. The probe pressure at one end, the predicted and measured pressure at the other end of the evaluation tube.

2.2.3 Prediction in real ear

The same approach was applied to the real ears of 5 human subjects to measure the ear-canal impedance and to predict the sound pressure at the TM. One example of the ear-canal impedance Z_{ear} is shown in Figure 8 (subject # 1). It was observed that the ear-canal impedance Z_{ear} resembled to that of the evaluation tube (Figure 6). The amplitude of the impedance $\left|Z_{ear}\right|$ had its first notch at about 3.3 kHz (f_0). A second notch of $\left|Z_{ear}\right|$ was present at 9.18 kHz, which was close to the third harmonic of f_0 . A peak was found between the two

notches at the frequency around $2f_0$. Note that the notches were not as deep as the impedance of the evaluation tube. The phase $\angle Z_{ear}$ showed more fluctuations and more gradual transitions at the notch or peak frequencies, compared with the tube impedance. The ear-canal impedance in the other 4 subjects showed similar patterns. Of course, the notch (or peak) frequencies differed among subjects, depending mainly on the length of the ear canal and the insertion depth of the eartip. The notch (or peak) shifted to higher frequency if the ear canal was shorter with a deeper insertion.
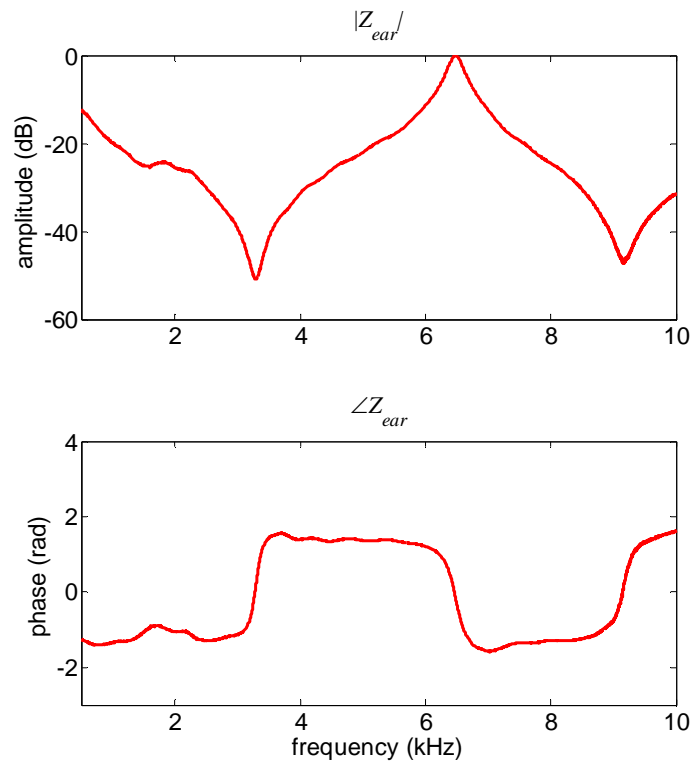
Figure 8. The amplitude and phase of the ear-canal impedance (subject # 1).

The estimated ear-canal impedance was then used to predict the TM pressure in all the 5 subjects and the predictions were validated by a second

microphone with a probe tube inserted approaching the TM. An example of the predicted TM pressure ($\hat{P}_{TM}$), the actual TM pressure (P_{TM}) measured by the second microphone and the probe pressure measured at the entry of the ear canal (P_L) is shown in Figure 9 (subject # 1). Generally, the observations in the ear canal were similar to those in the evaluation tube (Figure 7). The P_L and P_{TM} (or $\hat{P}_{TM}$) showed the greatest difference (up to 15 dB) around $f_0 = 3.3$ kHz and its third harmonic. Compared with the tube, the notches of P_L in the ear canal were not as deep. Around the frequency of $2f_0$, all measurements showed a peak. The prediction of the TM pressure was also successful in the ear canal, as indicated by the overall match between the $\hat{P}_{TM}$ and the P_{TM}. The mean prediction error averaged across frequencies was 0.8 dB and the standard deviation was 0.5 dB. The maximal prediction error was 2.5 dB, which was slightly larger compared with the prediction in the tube.
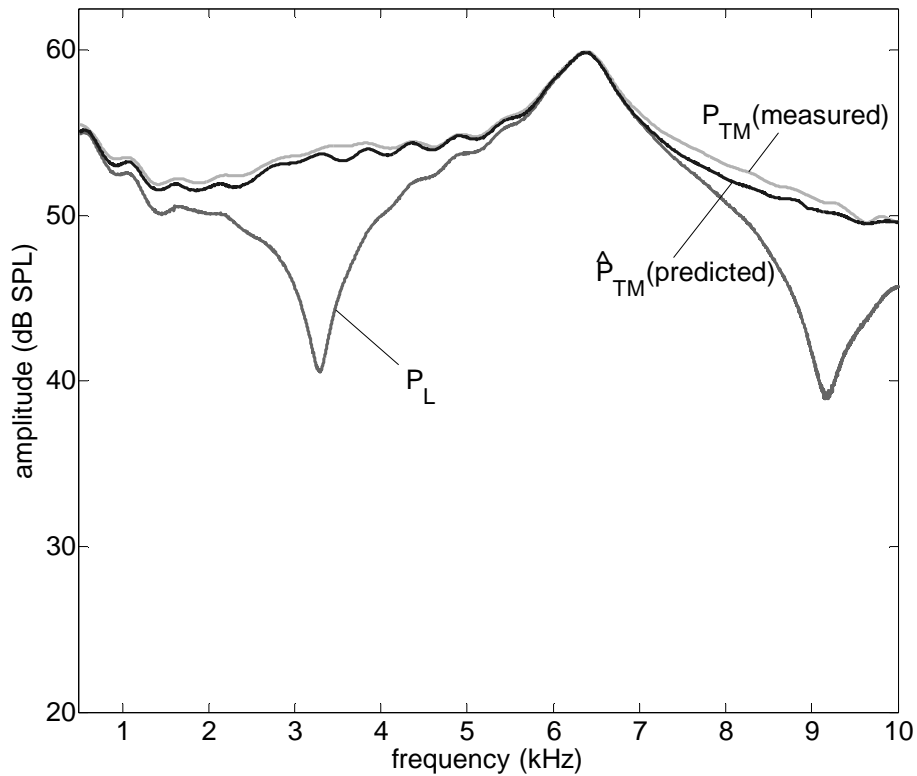
Figure 9. The probe pressure, predicted and measured TM pressure in the real ear measurement (subject # 1).

The prediction errors ($P_{TM} - \hat{P}_{TM}$) were calculated for all frequencies from 0.5 to 10 kHz in all the 5 subjects. The occurrences (or frequency counts) of various differences were expressed by a histogram for each subject (Figure 10). The range of the prediction error was from -2 to 3 dB among all the subjects, with an average ranging from 0.3 to 0.8 dB. These findings indicated that the actual sound pressure at the TM generally agreed with the predicted TM pressure in all the subjects. In other words, the sound pressure at the TM could be reliably estimated from the sound pressure at the entry of the ear canal based on the source calibrations in multiple cavities.
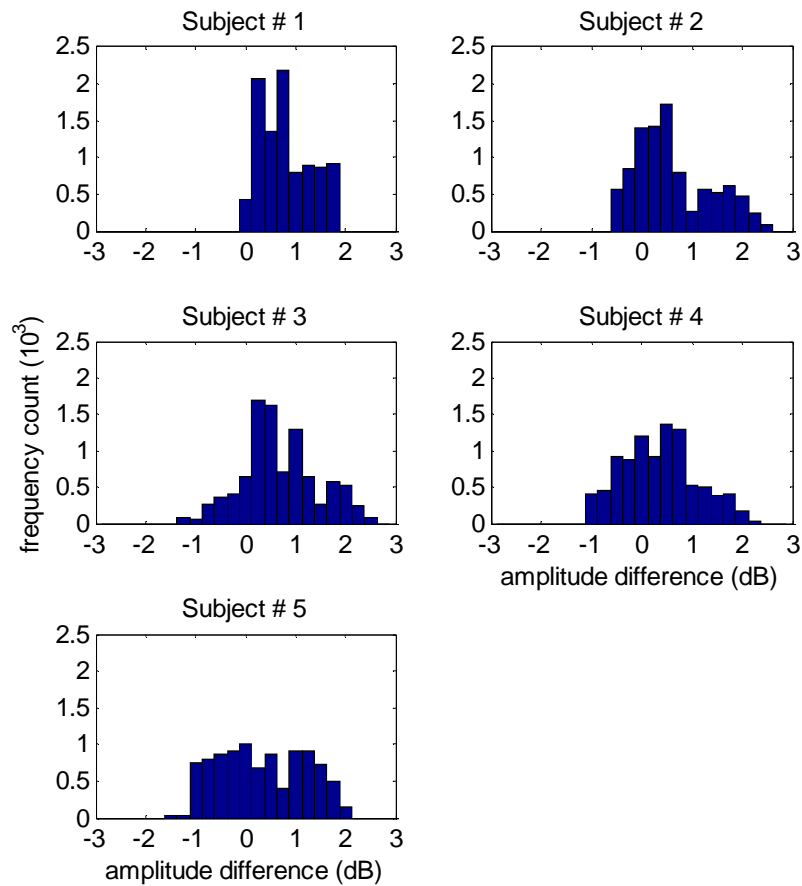
Figure 10. The histograms of the amplitude difference between the predicted and measured TM pressure (prediction error) of all subjects (n=5).

2.2.4 Effects of calibration methods on OAE measurements

There were 3 signal calibration approaches for OAE measurements: namely no calibration or compensation based on ear-canal sound pressure, calibration or equalization of sound energy at the ear-canal entrance, and at the TM. Effects of the three types of calibration strategies were evaluated by measuring swept-tone evoked SFOAEs (Figure 11). Panel (A) showed the probe pressure (P_L) and the corresponding predicted TM pressure ($\hat{P}_{TM}$) in the no

calibration condition without compensation. Similar to previous observations, the P_L had a deep notch at around 3.5 kHz due to the standing wave in the ear canal, while the $\hat{P}_{TM}$ was relatively flat. Both P_L and $\hat{P}_{TM}$ had a peak around 8.5 kHz.

As observed in panels (B)-(D), the swept-tone SFOAEs amplitude-frequency functions demonstrated two features: the global shape (baseline amplitude) and the fine structures (ripples). Since the fine-structures were generally similar across conditions, the effects of the calibrations were demonstrated by the changes in the global shape of OAE amplitudes. For the no calibration condition, the OAEs around 8.5 kHz showed an elevation of about 7 dB compared with other methods [noted by rectangle (1)], since the sound pressure entering the middle ear was nearly 10 dB higher than the rest of the spectrum. For the probe calibration condition, large compensations (up to 12 dB) were added around the notch frequency to equalize the sound energy at the entry of the ear canal. As a consequence, the sound energy entering the middle ear would be much higher than expected, and the OAE amplitudes around the notch frequency were elevated by as much as 6 dB [indicated by rectangle (2)]. In contrast, the OAE amplitude with the in-situ calibration did not show these undesired elevations since the stimulus level was equalized at the TM [panel (D)].

The effects of calibrations methods on OAE amplitude were similar in the other 4 subjects except that the frequencies of the notch and peak differed from subject to subject, depending on the effective length of the ear canal. The averaged elevation of the OAE amplitude across subject was 5.2 dB around the peak frequency for the no calibration condition, and it was 4.3 dB around the

36

notch frequency for the probe calibration condition. These observations suggested that the SFOAEs measured under the in-situ calibration condition did not suffer from the problem of inaccurate spectral amplitudes of the stimulus.
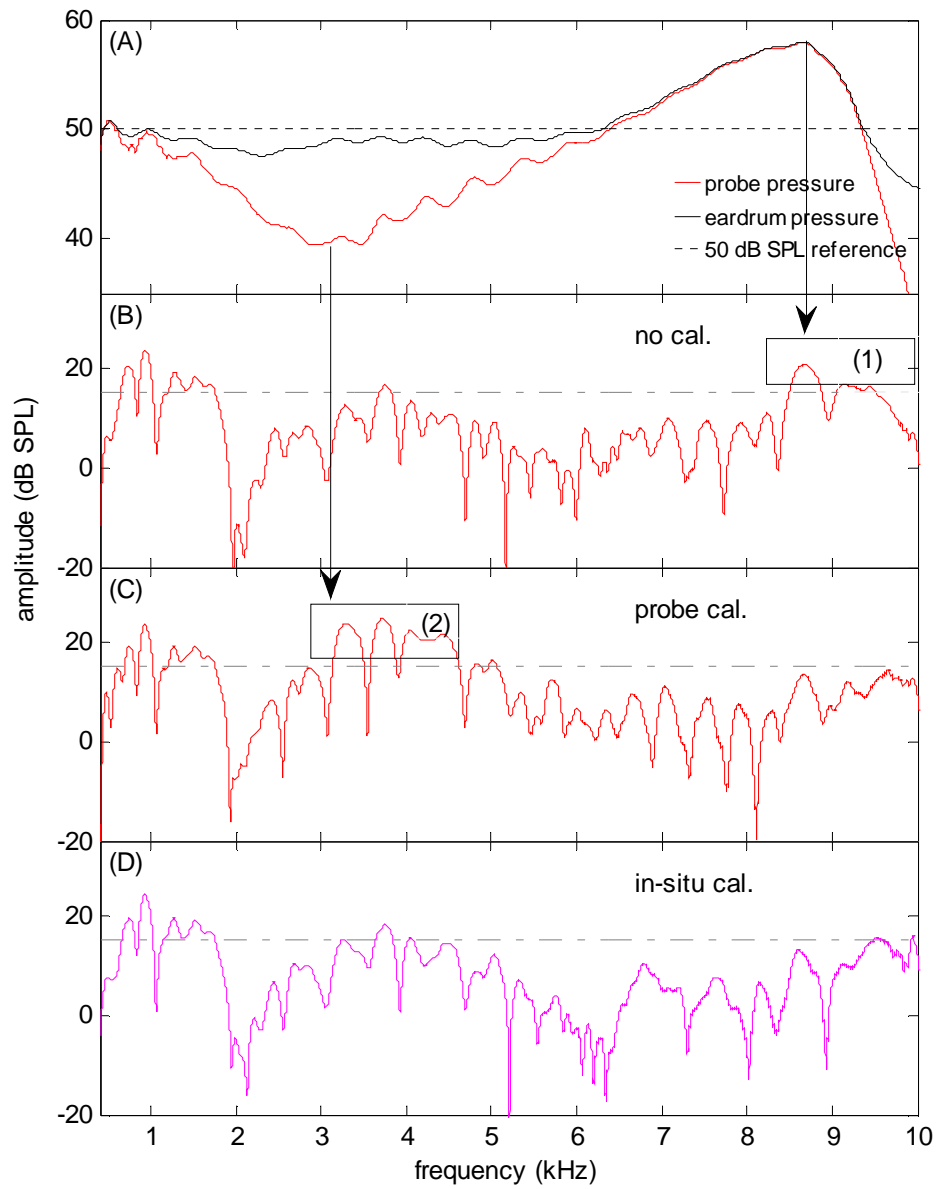
Figure 11. The effects of three different calibrations (no calibration, probe calibration and in-situ calibration) on the amplitude of swept-tone SFOAEs (subject # 2).

2.3 Discussion

2.3.1 Standing waves and ear-canal acoustics

In this study, it was found that there were notches and peaks in the sound

pressure measured at the entry of the ear canal or the tube (Figure 7 and 9),

although the voltage presented to the earphone was constant. These variations in

the probe pressure can be attributed to the standing wave, a well-known

phenomenon in the ear canal (Whitehead, Stagner, et al., 1995; Siegel, 1994;

Gilman & Dirks, 1985). The standing wave is formed by the interactions of the

incident and the reflected sound pressure waves. The interactions are dependent

on the phase relation between the two components. Since the reflected sound

wave travelled a distance of twice the length of the ear canal (L) to mix with the

incident pressure at the entry of the ear canal, the two components will be 180° out

of phase and cancel each other if distance traveled equals $1/2$ of the wavelength

of the signal ($2L = \lambda/2$), i.e., L = $1/4$ λ. Similarly, the cancellation will happen

when L equals $n/4 \cdot \lambda$ (n=3, 5, ∟). This can explain the presence of the notch in

the probe pressure at the quarter-wavelength frequency and its odd-order

harmonics. However, if the L is $n/2$ ($n = 1, 2, 3$∟) of the wavelength, the

reflection wave is in-phase with the incident pressure and the two components can

interact constructively, resulting a pressure peak at the half-wavelength frequency.

However, such a cancellation and enhancement do not occur at the closed-end of

the evaluation tube or at the TM where the two components are always in phase.

Therefore, the reflection wave can account for the large sound pressure difference

between the entry of the ear canal and the TM, especially at the quarter-wavelength frequency.

The standing wave problem exists in both the evaluation tube and the ear canals, and some of the observations are similar, because of their similarity in physical shapes. However, there are also differences between the results of the tube and the ear canal. The pressure notches can be as deep as 50 dB in the tube, and only 15 dB in the ear canals. This is probably because the sound wave can be completely reflected by the solid wall at the end of the tube, while only partial sound energy is reflected back at the TM, resulting in a partial cancellation in the real ear. Other factors, such as a slight curvature and differences in cross-sectional area, may also contribute to the differences in the observations between the tube and the ear canal. Furthermore, individual differences in the acoustic characteristics of the ear canal may exist between male and female, children and adults, and even different subjects within each group (Kruger & Ruben, 1987). Therefore, the acoustic characteristics of the ear canal must be measured for each individual prior to OAE measurements for more accurate signal calibration (Dalmont, 2001; Stevens et al., 1987).

2.3.2 Source calibrations

The source calibration is a fundamental step when predicting the TM pressure. The accuracy of the source calibration can be affected by many factors, such as the number of the tube lengths selected (the number of piston positions in this study), the diameter of the calculation tube and the frequency response of the earphone (Allen, 1986). Theoretically, two tube lengths should be sufficient since

only two unknown variables need to be solved (P_s and Z_s). However, the solutions may be inaccurate around the notch frequencies where the sound pressure measurement is not accurate, and increasing the number of measurements with different tube lengths can overcome this problem. Studies (Neely & Gorga, 1998; Keefe et al., 1992) recommend that 4 to 6 tube lengths are enough. Since the results from 5 and 6 tubes are similar with limited improvement, 5 different tube lengths were used in the present study. Regarding to the tube diameter, we used the same calculation tube with a diameter close to the averaged adult ear-canal since the participants were all adults. For measurement in children, different diameters of calculation tubes should be chosen. The frequency response of the earphone can influence the source calibration because the pressure measurement may be inaccurate if the output pressure from the ER-2A earphone is not large enough at high frequencies. This may explain the source parameters were not as accurate at high frequencies above 8 kHz (Figure 5). Scheperle et al (2008) found that the temperature is also a factor and they suggested that the calibration tube should be heated to body temperature.

When the results of the source calibration were used to estimate the impedance of the evaluation tube, the estimated impedance matched the theoretical impedance well (Figure 6). Since the accuracy of the impedance estimation is highly dependent on the reliability of the source parameters, the well matched impedances indicate that the source parameters obtained from the source calibrations are reliable and accurate.

2.3.3 Prediction of TM pressure

A key observation of this study is that the sound pressure at the end of the tube or the TM can be accurately predicted from the measurement at the entrance (Figure 7 and 9). Over the frequency range of 0.5-10 kHz, the maximal prediction error was 1.2 dB in the evaluation tube and less than 3 dB in the ear canals of all the 5 participants. The prediction was best around the peak frequency, and the prediction error seemed to increase at the notch frequencies and high frequencies above 8 kHz. Since the predicted TM pressure is calculated from P_L, P_s and Z_s, it is affected by the accuracy of the measurement of P_L. At the notch frequency, the amplitude of P_L is low and vulnerable to background noises, leading to larger prediction errors. For similar reasons, the limited output of sound pressure from the earphone may be responsible for the increased prediction error at high frequencies.

The advantage of predicting the TM pressure is that it does not require a direct measurement near the eardrum. Although direct measurement is an ideal way to obtain the actual pressure at the TM (Siegel, 1994; Siegel & Hirohata, 1994), deep insertion of a probe microphone could lead to discomfort of the subjects, especially for young children who may not be able to tolerate the approach. Another advantage is that the predicted TM sound pressure can be verified by a second microphone. This is different from other similar methods where the forward (incident) sound pressure is predicted (Lewis et al., 2009; McCreery et al., 2009; Scheperle et al., 2008; Johnson et al., 2007; Neely & Gorga, 1998). Although the forward pressure is also not affected by the standing

wave, it cannot be actually measured anywhere in the ear canal and therefore its

reliability cannot be validated. Predicting the sound pressure at the TM in the

present study permits the assessment of its validity.

One disadvantage of the TM pressure prediction, as well as all other

methods based on the multiple-cavity calibration, is that it requires additional time

(about 5 min) to recalculate the parameters of the earphone source before each

ear-canal measurement to produce the best prediction. An alternative solution is

that a single source calibration is used for measurements of all subjects. Neely and

Gorga (1998) tested the alternative solution to measure the behavior thresholds

and found that the threshold changes were less than 1 dB. Their results suggested

that even with single source calibration, the prediction method retains high

performance.

2.3.4 Effects of calibrations on OAE measurement

In this study, it is observed that different calibrations can impose different

effects on the swept-tone SFOAE amplitudes (Figure 11). The effects are

summarized as the following: 1) in the no calibration condition, the peak of the

stimulus amplitude at the half-wavelength frequency caused significant elevation

of the OAE amplitude; 2) in the probe calibration condition, the stimulus level

was over-compensated around the notch frequency and it caused considerable

elevation of the OAE amplitude; 3) in the in-situ calibration, the OAE amplitude

is more consistent across frequencies, and no undesired elevations of OAE

amplitude were found. A straightforward explanation is that the OAE amplitude

should be elevated accordingly if the TM pressure (or sound energy entering the

middle ear) is higher than what it should be. The most appropriate approach is the in-situ calibration which equalizes the sound pressure at the TM and brings the sound pressure to the desired level so that unexpected elevations of the OAE amplitudes can be avoided.

Since the stimulus is equalized across frequency at the TM, the OAE spectrum with the in-situ calibration is more reliable for possible use for comparing the OHC functions of different frequency regions. It is observed from Figure 11 (D) that the global shape of the OAE amplitude shows apparent reductions at about 2 kHz and 6 kHz. One possible explanation is that the OHCs may be less active, or have some slight damage, at the cochlear region of 2 kHz and 6 kHz for this subject. Similarly, the relatively large amplitude of OAEs at 1 and 4 kHz may be explained by the more active OHC functions at these frequency regions. However, other factors, such as the transfer function of the middle ear (Puria, 2003; Paul Avan et al., 2000; Osterhammel et al., 1993), should be taken into account before this method can be used for clinical diagnosis of the OHCs functions. Therefore, the normative spectral shape of the SFOAEs should be established with a population study from normal human beings.

The effects of different calibrations on the measurement of OAEs (mainly the DPOAEs) were also investigated by others. Siegel and Hirohata (1994) calibrated the stimulus level according to the probe pressure and measured eardrum pressure, respectively. They found that the difference between the measured DPOAE amplitudes of the two calibrations could be as large as 20 dB. Scheperle et al (2008) manipulated the probe-insertion depth during DPOAE

measurement and found that the results from the calibration according to the sound pressure at the entry of the ear canal demonstrated the greatest variability. Burke et al (2010) calibrated the stimulus level according to the probe pressure and forward pressure respectively. Surprisingly, no effects of the calibrations on the DPOAE amplitude were found. One reason for their negative finding may be that the frequencies they measured are too sparse (they measured DPOAEs at only five frequencies: 2, 3, 4, 6 and 8 kHz). The notch of the probe pressure may not be present at or close to these frequencies, and therefore the over-compensation problem was not maximally observed. In contrast, using the swept tone SFOAEs provides a chance to closely monitor the change of OAE levels with a high frequency-resolution.

2.3.5 Clinical implications

This study showed that the common calibration method that equalizes the sound pressure at the entry of the ear canal can cause over compensations of the stimulus level because of the existence of the standing waves. Moreover, moving the probe microphone close to the TM is not feasible for routine use, predicting the sound pressure at the TM or the in-situ calibration becomes an attractive technique to precisely control the sound energy entering the middle ear. The precise prediction of the TM pressure depends on reliable source calibrations prior to ear-canal measurements. Among other calibration methods, the in-situ calibration based on the predicted TM pressure yields OAEs measurements that are the least susceptible to adverse effects of inaccuracy in signal calibration.

45

Therefore, it is recommended by this study that the in-situ calibration method

should be further studied and be used in clinical OAE measurements.

Chapter 3

EXPERIMENT II

Otoacoustic emissions (OAEs) are soft sounds that are generated from the

inner ear and the normal activities of the cochlear outer hair cells (OHCs).

Therefore, they are widely used as a clinical tool to assess the inner ear function

in the clinic. Currently, two types of OAEs are measured for hearing screen and

diagnosis: transient evoked otoacoustic emissions (TEOAEs) and distortion

product otoacoustic emissions (DPOAEs). TEOAEs are evoked by clicks and

present in almost all ears with normal hearing, as long as the outer and middle

ears are normal (Probst et al., 1991). Clicks are short duration and yet broadband

stimuli that can stimulate the entire basilar membrane simultaneously. Thus

TEOAE test can measure a wide range of frequencies quickly, making it an ideal

screen tool. DPOAEs are induced by two pure tones and extra sound energies are

present at frequencies other than the frequencies of the two tone stimuli. DPOAEs

can be extracted by performing a simple spectral analysis of the ear-canal acoustic

signal. DPOAEs are good indicators of the cochlear nonlinearities due to OHC

transductions (Bian & Scherrer, 2007; Bian, 2004; Bian et al., 2002). They are

useful in detecting and monitoring hearing functions within a specific cochlear

region (Davis et al., 2004; Seixas et al., 2004; Attias et al., 2001; Telischi et al.,

1995). The input-output functions of DPOAEs have been recently used to assess

cochlear gain control (Chen & Bian, 2011).

However, both TEOAEs and DPOAEs have limitations that restrict their

clinical applications to mostly hearing screening, instead of comprehensive

assessment of the inner ear function and differential diagnosis of auditory disorders. Although TEOAEs are generally believed to be generated from linear reflections of the traveling waves by the OHCs at the CF place of each frequency component (Shera & Guinan, 1999). However, there are uncertainties of the emissions generated at each frequency component. Time-frequency analysis of TEOAE temporal waveform showed that an individual frequency component could occur at multiple latencies, suggesting that multiple reflections may exist to complicate the interpretation of the TEOAE results (Goodman et al., 2009; Hatzopoulos et al., 2000). Distributions of acoustic energy at a particular time moment can spread to a range of frequencies, especially at high frequencies (Tognola et al., 1997). Avan et al (1993) found that the cochlear damages of high frequencies (6-8 kHz) could lead to the reduction of TEOAE amplitude at low frequencies (1 kHz), suggesting that the TEOAE component at an individual frequency may consist of energy from multiple cochlear regions. Moreover, the stimulus artifacts which occur at the beginning of the recording, thus the first 2-3 ms of the recorded sound waveform is removed from analysis and therefore the high-frequency OAEs (above 4 kHz) with short latencies are absent in TEOAEs (Kemp et al., 1990). DPOAEs are not affected by stimulus artifacts and they can be measured from high frequencies. However, their generation mechanisms are complicated. They are considered to be a mixture of two components generated from two distinct cochlear locations via different mechanisms: a nonlinear distortion component from the overlapped region of the two travelling waves and a linear reflection component from the CF place of the DP frequency (Shera &

Guinan, 1999). The complicated generation mechanisms make it difficult to interpret DPOAE results or to relate DPOAE abnormities to a specific location of cochlear damages. Other investigators (Long et al., 2008; Kalluri & Shera, 2001) managed to unmix the two components, however, it involves another level of data analysis and thus another source of variability. In addition, the exact location of the distortion component still remains unclear because the overlapped region includes an extended area basal to the f_2 place (G. K. Martin et al., 2010; Wilson & Lutman, 2006). Furthermore, DPOAEs are measured frequency by frequency, making it less efficient compared with TEOAEs. Due to the above limitations, TEOAEs and DPOAEs are mainly used to determine whether the cochlea is normal or impaired, and detailed information about cochlear functions is often missing.

To further characterize the cochlear OHC functions with more accurate frequency specificity, a third type of OAEs, the stimulus frequency otoacoustic emissions (SFOAEs), show some unique strengths that could be useful to overcome the limitations of TEOAEs and DPOAEs. SFOAEs are emissions evoked theoretically by one single tone. Although they were discovered along with other types of OAEs in 1978, SFOAEs are currently not used in clinical practices. One reason for the lack of clinical use is that the SFOAEs have the same frequency as the evoking stimulus, making it complicated to separate them from the stimulus. However, SFOAEs have several distinct advantages that weigh over the limitations of other OAE measurements. First, SFOAEs are considered to be generated from the reflection within a localized region of the peak of the

traveling wave of the stimulus (Cheatham et al., 2011; Talmadge et al., 2000; Kemp & Chum, 1980). Such generation mechanism is simpler than TEOAEs or DPOAEs that involve multiple generation regions. Therefore, SFOAEs could be sensitive to small damages of the generating OHCs and the abnormality in SFOAEs could map out specific regions of dysfunctional OHCs. Second, the amplitude of SFOAEs is dependent on the amplitude of the forward travelling wave, which is in turn determined by the nonlinearities of the normal OHCs. Therefore, the measurement of SFOAEs can reflect various aspects of the OHC nonlinearities, such as the compressive growth (Schairer et al., 2003) and the two-tone suppression (Kemp & Chum, 1980). Third, studies (Siegel et al., 2005; Shera et al., 2002) showed that the phase-frequency function of SFOAEs could be used to estimate the group delay of the basilar membrane (BM) and the change of the BM group delay could be an indicator of various cochlear damages. Fourth, the reflected and forward travelling waves can interact with each other in the pathway, causing frequency-dependent cancellations or enhancements (Goodman et al., 2003; Zweig & Shera, 1995). Such cancellation or enhancement gives rise to the well-known phenomenon of SFOAE amplitude: fine structures (quasi-periodic amplitude fluctuations across frequencies). Fine structures are great indicators of normal OHC activities (Rao & Long, 2011; Wagner, Plinkert, et al., 2008; Talmadge et al., 1998) and they are correlated with the microstructures of the hearing threshold (Lutman & Deeks, 1999).

Two special techniques are used to extract the SFOAEs since the ear-canal signal contains two components (the stimulus and the SFOAEs) which have the

50

same frequency. One approach, the compression method, makes use of the compressive growth of SFOAE amplitude and removes the stimulus component by scaling down the ear-canal response to a higher level of stimulus to obtain a nonlinear residue as the SOFAE component (Kemp, 2002; Kemp et al., 1990). Another approach, the suppression method, tries to measure the ear-canal signal with and without a suppressor tone in addition to a probe tone. Because the SFOAEs can be suppressed by the second tone, the difference between the ear-canal signals under the two conditions is recorded as the SFOAEs (Kalluri & Shera, 2007a; Kemp & Chum, 1980). Regardless of the different methods, the extracted SFOAEs are nearly equivalent (Kalluri & Shera, 2007a; Schairer & Keefe, 2005), which provides strong support for possible clinical applications of SFOAEs.

A common feature of current SFOAE extraction methods is that they all use pure tones as stimuli. Since SFOAEs evoked by pure tones need to be measured frequency-by-frequency, very limited number of frequencies can be measured and the SFOAEs are usually measured within a relatively narrow frequency range. A solution is to use a frequency-sweeping signal as the stimulus so that multiple frequencies of SFOAEs can be measured within one single test. There are two types of frequency-sweeping signals commonly used: chirps and swept tones. Chirps are short duration swept-frequency sinusoids commonly used in radar (Roos et al., 2009; Levy & Azhari, 2007). Chirps have freely selectable frequency range, and they can stimulate selected range of cochlear partitions to evoke emissions. However, the emissions evoked by chirps are similar to

51

TEOAEs evoked by clicks (Fobel & Dau, 2004; Keefe, 1998; Neumann et al., 1994). This is because the chirp is a type of transient signal in which multiple frequencies occur within a rather short period of time. Emissions of different frequencies are not well separated in time and multiple reflections can occur to complicate the chirp-evoked OAEs. Swept tones are frequency-sweeping signals originally designed for the measurement of the transfer function of the audio system, and the use of swept tones shows significantly higher immunity against system distortion (Müller, 2001). Swept tones also have customizable frequency range, allowing for the functional evaluation of a selected cochlear region. Longer durations in swept tones can ensure that there is enough dwell time for each frequency, so that OAEs are measured in a similar way as pure-tone test. As expected, a high resolution OAE gram (OAE amplitude as function of frequency) with the spectrum free from the contaminations of multiple reflections could be obtained. However, using swept tones to evoke emissions has not been reported.

In this study, attempts have been made to measure SFOAEs evoked by swept tones. An approach for extracting swept-tone SFOAEs was developed, and the effects of different signal conditions on the amplitude and phase of swept-tone SFOAEs were explored. The test-retest reliability of the swept-tone SFOAEs was investigated by a repeated-measure design, and the accuracy was examined by comparing the results with a standard approach using pure tones. The in-situ calibration described in experiment I was used to calibrate the level of the swept-tone stimulus.

3.1 Method

3.1.1 Subjects

Twenty subjects (5 males and 15 females) with ages ranging from 23 to 33 years old (mean age=25.9) participated in this experiment. The subjects had normal hearing with thresholds of 20 dB HL or less for frequencies from 250 to 8000 Hz in behavioral tests. No history of outer or middle ear surgery was reported from the subjects. An otoscopic examination was performed prior to the test and excessive earwax was removed to ensure a clean ear canal. Before the test, both of the ears of each subject were briefly screened using the swept-tone SFOAE program and only the ear with larger overall amplitude of OAEs was measured. The subjects were instructed to be as quiet as possible during the test. The subjects were seated in a sound-proofed booth comfortably during the test. Quiet reading and sleep was allowed during the test because no attention or behavioral response was required for the OAE measurements. The recruiting protocols were approved by the Institutional Review Board of ASU.

3.1.2 Equipment

A custom program was developed in Labview (National Instruments, Austin, TX) to generate the stimuli from the computer. The generated signals were delivered to two ER-2A earphone probes (Etymotic Research, Elk Grove Village, IL) through two channels on the PXI-4461 card (National Instruments). The earphones were seated together with the ER-10B+ probe microphone (Etymotic Research) in the eartip whose size was selected according to the ear-canal size of each subject. The eartip was inserted into the ear canals to deliver the

acoustic stimuli and the sound pressure near the earphone was recorded by the probe microphone. The recorded response was then digitized by the PXI-4461 card at a sample rate of 80k sample/s.

3.1.3 Stimuli

A) Stimulus presentation

A three-interval paradigm (Keefe, 1998) was used to measure the SFOAEs (Figure 12). Two ER-2A earphones (earphone A and B) were used to deliver two stimulus tones: the probe tone s_1 and the suppressor tone s_2. The s_1 and s_2 could be either the swept tones or pure tones. A ramping and fading of 0.02 ms were applied to the begging and end of s_1 and s_2 to avoid switching noises in the response. There were three intervals for presenting the stimuli in a particular manipulated order. In the 1st interval, earphone A played s_1 and earphone B played no sound. In the 2nd interval, earphone A played no sound and earphone B played s_2. In the 3rd interval, earphone A played s_1 while earphone B played s_2 simultaneously. The time waveforms of the acoustic responses of the three intervals were p_1, p_2 and p_{12}, respectively. Then the SFOAEs evoked by the swept tones could be extracted by the residual of the three responses:

$$p = p_1 + p_2 - p_{12} \qquad (5)$$

54
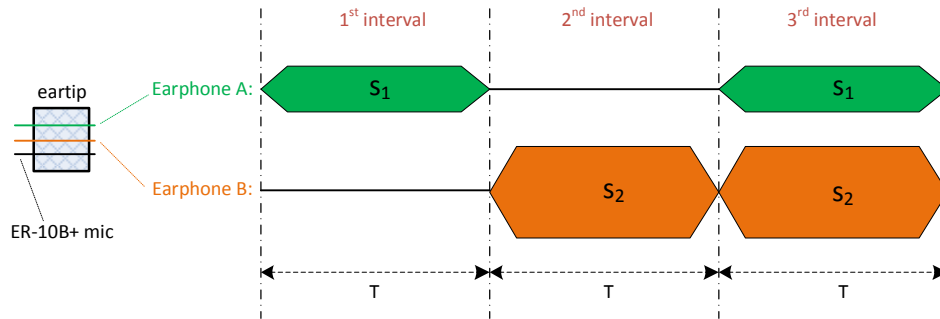
Figure 12. The three-interval paradigm to measure swept-tone SFOAEs.

This is because the stimuli (s_1 and s_2) were cancelled by the subtraction in equation (5). In contrast, the OAEs in the third interval (evoked by either s_1 or s_2) were smaller compared with the first two intervals, due to the two-tone suppression in the cochlea (Kemp & Chum, 1980). Therefore, the residual would contain a nonlinear portion of the SFOAEs evoked by both s_1 or s_2. The SFOAEs in this study referred to the SFOAEs evoked by s_1.

B) Swept-tone stimuli

To record swept-tone SFOAEs, s_1 and s_2 were both swept tones with linearly sweeping frequencies (Fig. 12). The frequency of s_1 increased linearly from 0.5 to 10 kHz within a duration of T which was changed from 0.25 to 2 s in the experiment. The s_2 was a similar swept tone whose frequency was always 200 Hz higher than s_1 (Figure 13) and the duration was the same as s_1. The amplitude of s_2 was kept constant at 80 dB SPL. The level of s_1 (L) was lower than s_2 and it was increased from 45 to 60 dB SPL.
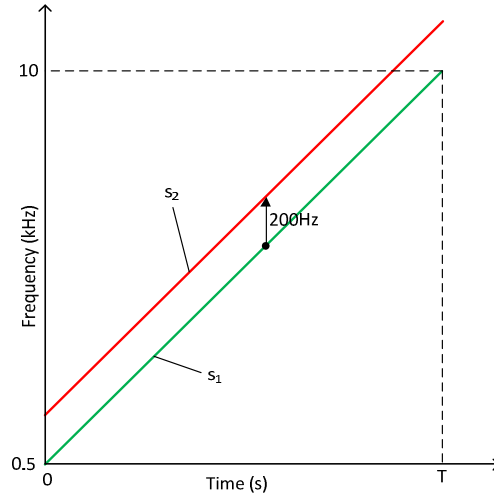
Figure 13. The frequencies of the swept-tone stimuli used to evoke SFOAEs.

3.1.4 Procedures

The source calibrations were performed on both earphones in multiple cavities to solve for the source parameters, following the procedures described in experiment I. Then the source parameters were used for the in-situ calibration of the level of the swept-tone or pure-tone stimulus.

Each subject participated in three successive sections in the experiment (Figure 14): two repeated sections of swept-tone SFOAEs (trial 1 and 2) and one section of pure-tone SFOAEs. The repeated measures of the swept-tone SFOAEs were to examine the test-retest reliability. The pure-tone SFOAEs served as a standard method to evaluate the consistency (or the accuracy) of the swept-tone SFOAEs. It took about 25 minutes for one trial of swept-tone SFOAE measurements, and 30 min for the pure-tone SFOAE measurement.

56

```
Swept-tone SFOAEs          Pure-tone SFOAEs          Swept-tone SFOAEs
   (Trial 1)          →                      →          (Trial 2)

    (25 min)                    (30 min)                    (25 min)
```
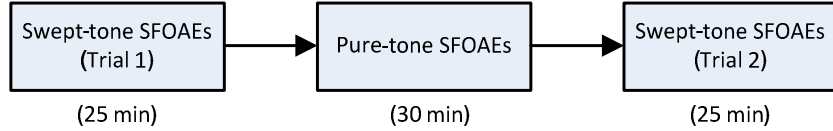
Figure 14. The three sections of successive measurements for the entire experiment.

In the first and third section, two swept tones (Figure 13) were used to evoke swept-tone SFOAEs. The level of swept tone s_1 (L) and the signal duration (T) were systemically varied to explore the effects of the signal conditions on the swept-tone SFOAEs. The L was increased from 45 to 60 dB SPL in a 5-dB step. The in-situ calibration was used to precisely control the L at the TM position. For each L, the swept tone was presented in four different durations: $T = 0.25,\ 0.5,\ 1,\ 2$ s. The acoustic response of each condition was averaged 30 times for each signal condition.

In the second section, the same three-interval paradigm was used to evoke pure-tone SFOAEs (Figure 12). Due to the time constraint, the probe tone s_1 was within a narrow frequency range from 1 to 2 kHz with 50 evenly distributed frequencies. The suppressor tone s_2 was 200 Hz higher than s_1. The level of the s_1 was set to 50 and 60 dB SPL successively, while the level of the s_2 was maintained at 80 dB SPL. The in-situ calibration was also used to set the level of the pure-tone stimulus. The duration of the both the s_1 and s_2 were 0.5 s. The response was averaged 10 times for each signal condition.

3.1.5 Data analysis

   To further extract the swept-tone SFOAEs from noises, a tracking filter
(Appendix E) was applied to the temporal waveform of the residual calculated
from Equation (4). The tracking filter was a time-varying narrow band-pass filter
which only passed the swept-tone SFOAEs evoked by s_1 and attenuated all other
irrelevant signals including the swept-tone SFOAEs evoked by s_2. Then a fast
Fourier transform (FFT) was applied to the filtered output to yield a spectral
complex which was further analyzed to obtain the amplitude and phase spectra of
the SFOAEs. A correction for the delay of the recording system was applied to
the phase of the spectral complex. The noise floor of the SFOAEs was obtained
by applying the same tracking filter to the residual with the center frequency set to
a frequency 200 Hz lower than the s_1 where there were no OAEs.

   To analyze the test-retest reliability of the swept-tone SFOAEs, the
amplitude and phase differences between results of the two repeated measures
were calculated for each of the 4×4 signal conditions in each subject. Then the
differences were averaged across frequency, and a two-way ANOVA was
performed to test whether there were significant effects of the stimulus level and
duration. The absolute amplitude difference and correlation coefficients were also
calculated to further examine the similarity of the amplitudes of the two repeated
measures.

   To extract the pure-tone SFOAEs, an FFT was applied to the residual of
each pure-tone measurement. The amplitude of the FFT output at the probe
frequency was taken as the amplitude of the SFOAE. The phase of the pure-tone

58

SFOAE was obtained by subtracting the phase of probe-tone stimulus from the phase of the FFT output. The spectral contents of the FFT output within 100Hz on both sides of the probe frequency were averaged to provide an estimate of the noise floor.

To analyze the consistency of the results between the two methods, the amplitude and phase of the swept-tone SFOAEs were sampled at the discrete probe frequencies used for the pure-tone SFOAEs. Then the amplitude and phase differences between the two methods were calculated at each frequency. The differences averaged across frequency in all subjects were analyzed by a t-test. The absolute amplitude difference and correlation coefficients were also calculated to further examine consistency between the two methods.

Specific features of the swept-tone SFOAEs were also analyzed in this study. The details of the analyzing methods were described in corresponding sections in the results.

3.2 Results

3.2.1 Overview

A) Validation of swept-tone SFOAEs

In order to verify the effectiveness of the three-interval paradigm in extracting the swept-tone SFOAEs, the same measurement was carried out in a tube (without OAEs) and an ear canal (with OAEs). The signal conditions were: the s_1 stimulus was 60 dB SPL and the frequency changed linearly from 0.5 to 10 kHz within 1 s. The s_2 stimulus was 80 dB SPL and its frequency was kept 200 Hz higher than the s_1. A tracking filter technique was used to extract the swept-

tone SFOAEs from the residuals calculated from equation (5). The temporal

waveform, amplitude and unwrapped phase spectra of the responses in the tube

and in the ear canal were compared in Figure 15. For the response in the tube, the

spectral amplitude never exceeded -15 dB SPL and the phase randomly fluctuated

around 0 across frequencies, indicating that the residual in the tube contained

mainly noises and no OAEs or stimulus-related residuals were found. In contrast,

the response in the ear canal was very different. The spectral amplitude of the

response could reach 20 dB SPL and apparent amplitude fluctuations (fine

structures) were observed. In addition, deep phase gradient was also observed

when the frequency increased. Since the fine structures and deep phase gradient

are unique features of SFOAEs (Goodman et al., 2003; Zweig & Shera, 1995), it

was proven that the three-interval paradigm could successfully obtain swept-tone

SFOAEs in healthy human ears.
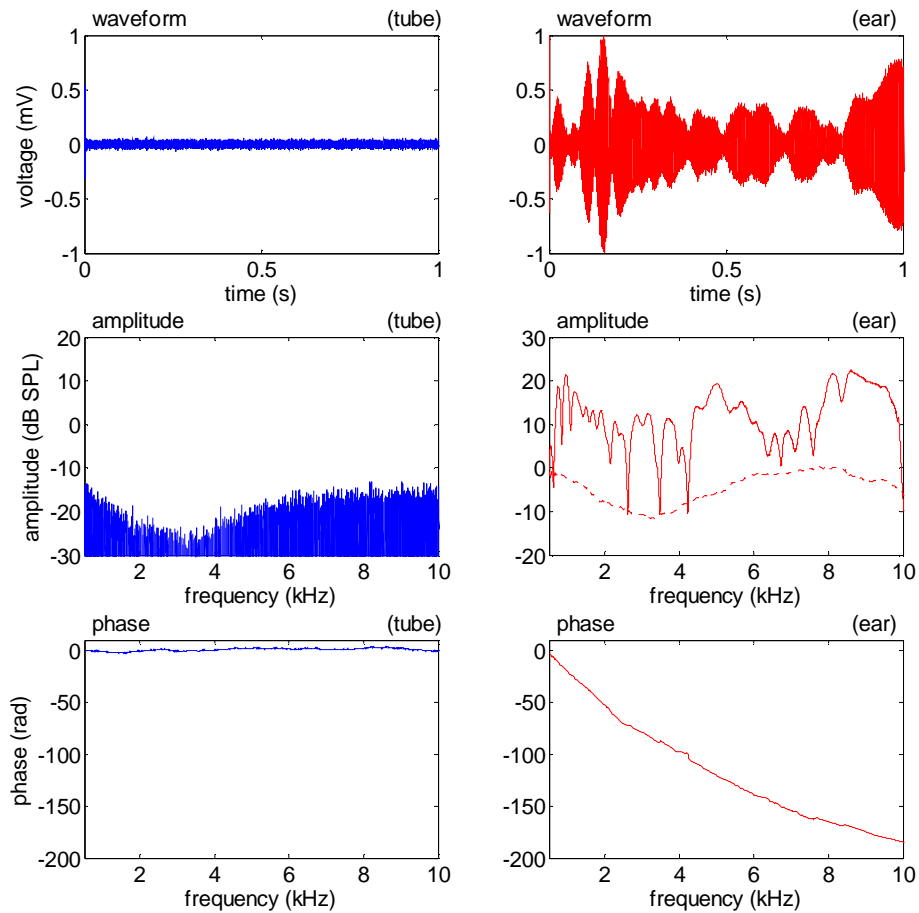
Figure 15. Temporal waveform, amplitude and phase spectra of the responses in the tube and in the ear (subject # 1).

B) General features

As noted in Figure 15, the amplitude and phase spectra of the swept-tone SFOAEs generated by human ears demonstrated general frequency-dependent features that made them differ from background noises or other artifacts. For the amplitude, it could be characterized in two aspects: the global shape (or the baseline) and the fine structures (or the ripples). Although the acoustic energy was equalized across frequencies at the TM, the global shape of the amplitude was not

61

constant with reductions presented at different frequencies regions. The fine structures were manifested by peaks and notches with regular frequency spacing, which seemed to be narrower for lower frequencies and wider for higher frequencies. The patterns of the global shape and fine structures were highly individual-dependent. Meanwhile, the phase of the swept-tone SFOAEs showed rapid reduction as the frequency increased. At low frequencies where the fine structures were denser, the phase dropped more rapidly. The amount of the phase reduction seemed to be proportional to the number of fine structures. There were some phase discontinuities at the frequencies where the amplitude showed deep notches.

3.2.2 Effects of signal conditions

In this experiment, swept-tone SFOAEs were measured in 20 normal-hearing subjects under 4 conditions of the level L and 4 conditions of the duration T. General features of the amplitude and phase (described above) were observed in all subjects under various signal conditions. In the following section, examples of the level and duration effects on these features were demonstrated, and then the results of all subjects were statically analyzed.

A) Level effects

An example of the amplitude and unwrapped phase of the swept-tone SFOAEs for four different stimulus levels were shown in Figure 16 ($T = 2$ s). For the amplitude, the global amplitude showed elevations as the stimulus level increased but the global shapes were similar. The fine structures of the amplitude became less pronounced and the frequency spacing seemed to be wider at higher

stimulus levels. Another observation was the apparent reduction of the slope of the phase-frequency function. If the total phase reduction was noted as θ within the frequency range from 0.5 to 10 kHz, it was noted that the θ reduced from 280 rad at 45 dB SPL to only 200 rad at 60 dB SPL. The level effects on the amplitude and phase of the swept-tone SFOAEs were similar for other signal durations.
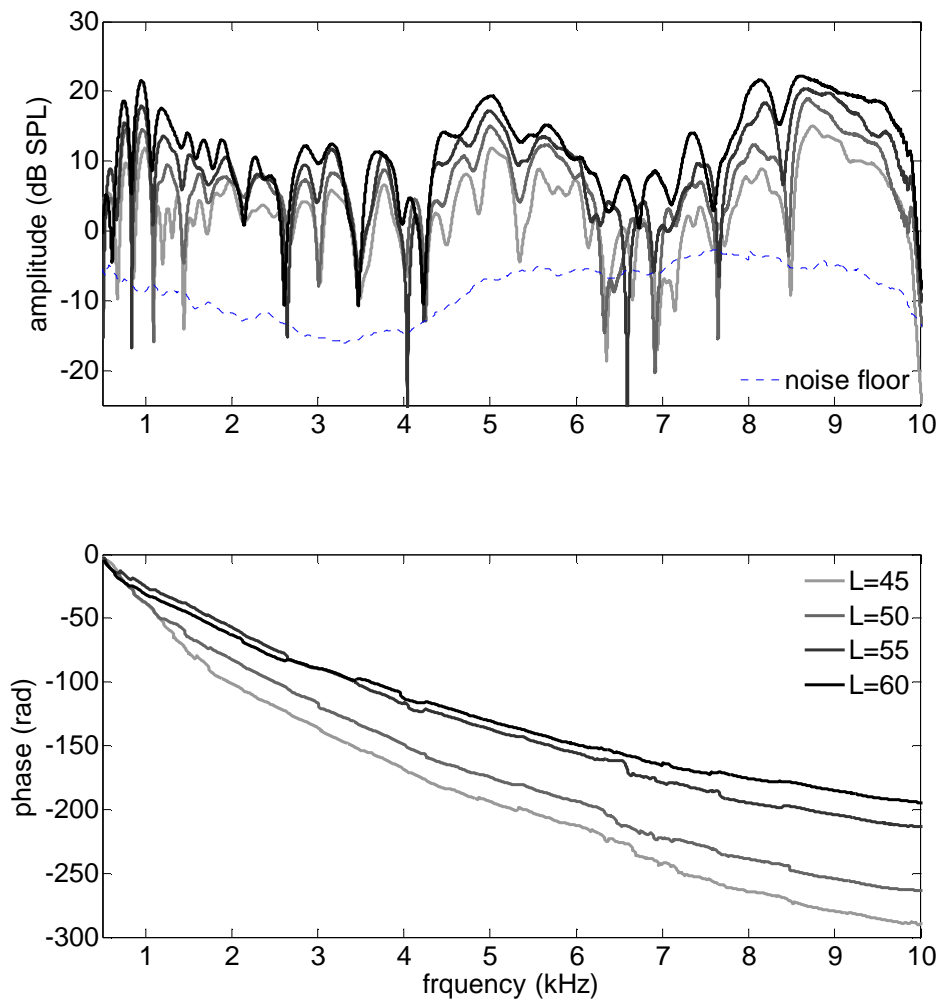
Figure 16. The amplitude and phase spectra of the swept-tone SFOAEs measured at four different stimulus levels (subject # 1).

B) Duration effects

An example of the amplitude and phase spectra of the swept-tone SFOAEs measured at 4 different stimulus durations was shown in Figure 17 ($L = 50\,\mathrm{dB}$ SPL). It was found that both the amplitude and phase spectra were very similar regardless of the nearly 10 fold of change in the duration (or sweeping speed). For the amplitude spectrum, both the global shape and the fine structures were not altered by the 4 different signal durations. There were some amplitude deviations at some frequency regions where the OAEs were smaller, but the changes were relatively random. For the phase spectrum, the curves corresponded to the four durations almost overlapped with each other. The deviation of the phase function of $T = 2$ was due to random phase jumps at notch frequencies. The duration also had no evident effects on the swept-tone SFOAEs measured under other stimulus levels.
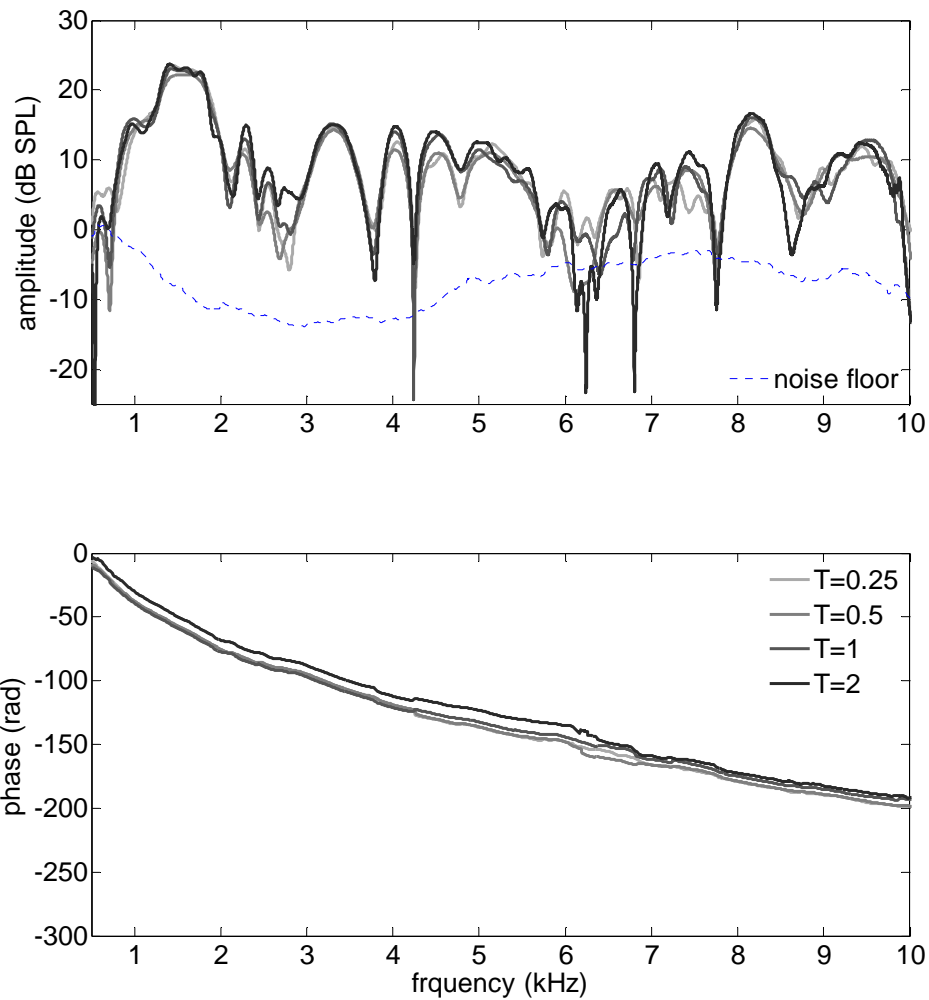
Figure 17. The amplitude and phase spectra of the swept-tone SFOAEs measured

with four different durations (subject # 2).

C) Signal effects across subjects

To confirm the signal effects among all subjects, the amplitude averaged

across frequencies from 0.5 to 10 kHz (A) and total phase reduction (θ) under

different signal conditions were shown in Figure 18. An ANOVA with two

independent variables (IVs: L and T) was performed on two dependent variables

(DVs: A and θ), respectively. The p-values for L were less than 0.001 for both DVs, indicating that L had significant effects on both DVs. However, no significant effects of T on either DV were found (p=0.30 for A, p=0.23 for θ). There were no significant interactions between the two IVs. These effects of the IVs could also be observed from Figure 18: as the L increased from 45 to 60 dB SPL, the averaged amplitude A increased from 5 to 10 dB SPL and the total phase reduction θ decreased from 220 to 150 rad; while for each L, the duration T did not seem to have significant effects on either A or θ because the error bars were overlapped.
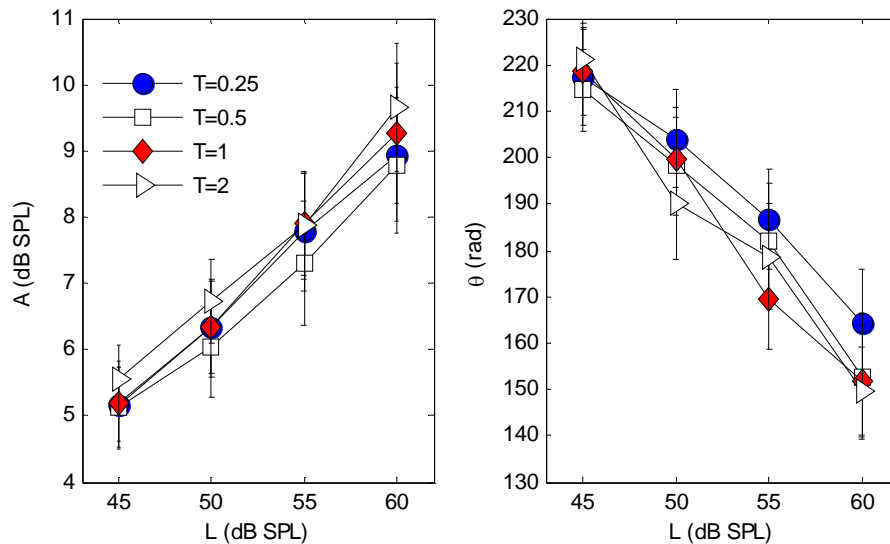
Figure 18. The mean and standard deviation of the amplitude and total phase reduction averaged across frequency (n=20).

3.2.3 Test-retest reliability

In this study, the test-retest reliability of the swept-tone SFOAEs was examined by comparing the results of two repeated measures (trial 1 and 2) with a

66

30 min interval. An example of the comparison was shown in Figure 19 (signal conditions: $L = 50, 60$ dB SPL; $T = 2$ s). It could be observed that the spectra of the amplitude and phase could be largely reproduced in the repeated measures, and the reproducibility was better at higher stimulus levels. For the amplitude spectrum [panel (A) and (B)], the two amplitude-frequency functions were very similar. The correlation coefficient (r) of the two functions was 0.79 for L=50 dB SPL and 0.90 for L=60 dB SPL. For the phase spectrum [(C) and (D)], the two phase-frequency functions from the two repeated measured greatly overlapped, except that the phase discontinuities were larger at the lower stimulus level. The difference between the SFOAE amplitudes of the two repeated measures was calculated at each frequency, and a histogram was constructed based on the difference [(E) and (F)]. The histograms were centered at 0 for both stimulus levels. The standard deviation was 3.6 dB for $L = 50$ dB SPL and 2.5 dB for $L = 60$ dB SPL, indicating that the averaged differences were essentially zero.
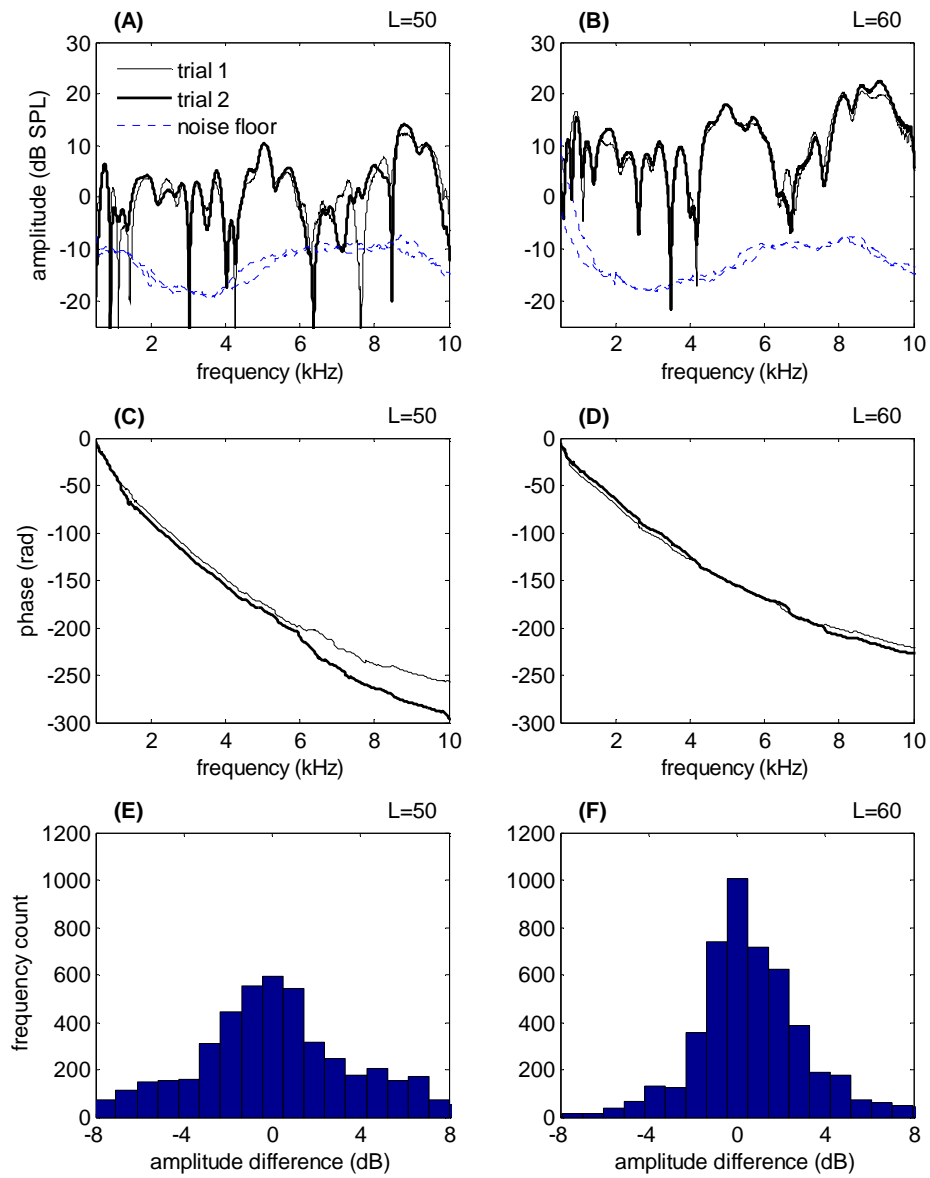
Figure 19. The amplitude and phase spectra of the swept-tone SFOAEs of two repeated measures (subject # 4).

For all subjects, the between-trial difference of the amplitude (ΔA) and the difference of the total phase reduction ($\Delta \theta$) were averaged across frequencies from 0.5 to 10 kHz. The mean and standard deviation of ΔA and $\Delta \theta$ under

different signal conditions were shown in Figure 20 (A)-(B). It was observed that the amplitude difference ΔA fluctuated around 0 and it never exceeded 1 dB. The phase difference $\Delta \theta$ varied randomly within -5 to 5 rad under different signal conditions. Then a two-way ANOVA (IVs: L and T) was performed on ΔA and $\Delta \theta$, respectively. No significant effects of the IVs were found, indicating that the mean of ΔA and $\Delta \theta$ did not significantly different from 0 under different signal conditions.

To further examine the similarity of the results of the two repeated measures, the absolute values of the amplitude difference averaged across frequency ($|\overline{\Delta A}|$) and the correlation coefficients of the two amplitude-frequency functions (r) were shown in Figure 20 (C)-(D). It was observed in panel (C) the $|\overline{\Delta A}|$ decrease from 2.5 to 2 dB as the stimulus level L increased from 45 to 60 dB SPL. For each L, the $|\overline{\Delta A}|$ of different durations were very similar. The correlation coefficients r in panel (D) increased within a range from 0.7 to 0.9 as either the L or the T increased. These findings indicated that the SFOAEs results were replicable and the increase either the stimulus level or the duration could reduce the between-trial deviation.

Figure 20. The amplitude difference (ΔA), phase difference ($\Delta\theta$), absolute amplitude difference ($|\Delta A|$) and the correlation of the amplitude-frequency functions (r) between the two repeated measures (n=20).

3.2.4 Consistency with pure-tone SFOAEs

The consistency with the standard methods using the pure tones represented the accuracy of the swept-tone SFOAEs. In this study, the results of the swept-tone SFOAEs were compared with the pure-tone SFOAEs measured under the same signal condition. An example of the results from the two methods

was shown in Figure 21. For the amplitude spectra (panels A and B), the two

amplitude spectra seemed to match better for large amplitude of SFOAEs and

more variable when the SFOAE amplitude was close to noise floor. For the phase

spectra (panels C and D), the curves from the two methods overlapped with each

other for most of the measured frequencies. The histograms in panel (E) and (F)

showed that the differences in SFOAE amplitude between the two methods were

centered at 0 dB with a standard deviation of about 1.5 dB. These observations

suggested that the SFOAEs from the two methods showed great consistency

under different stimulus levels.
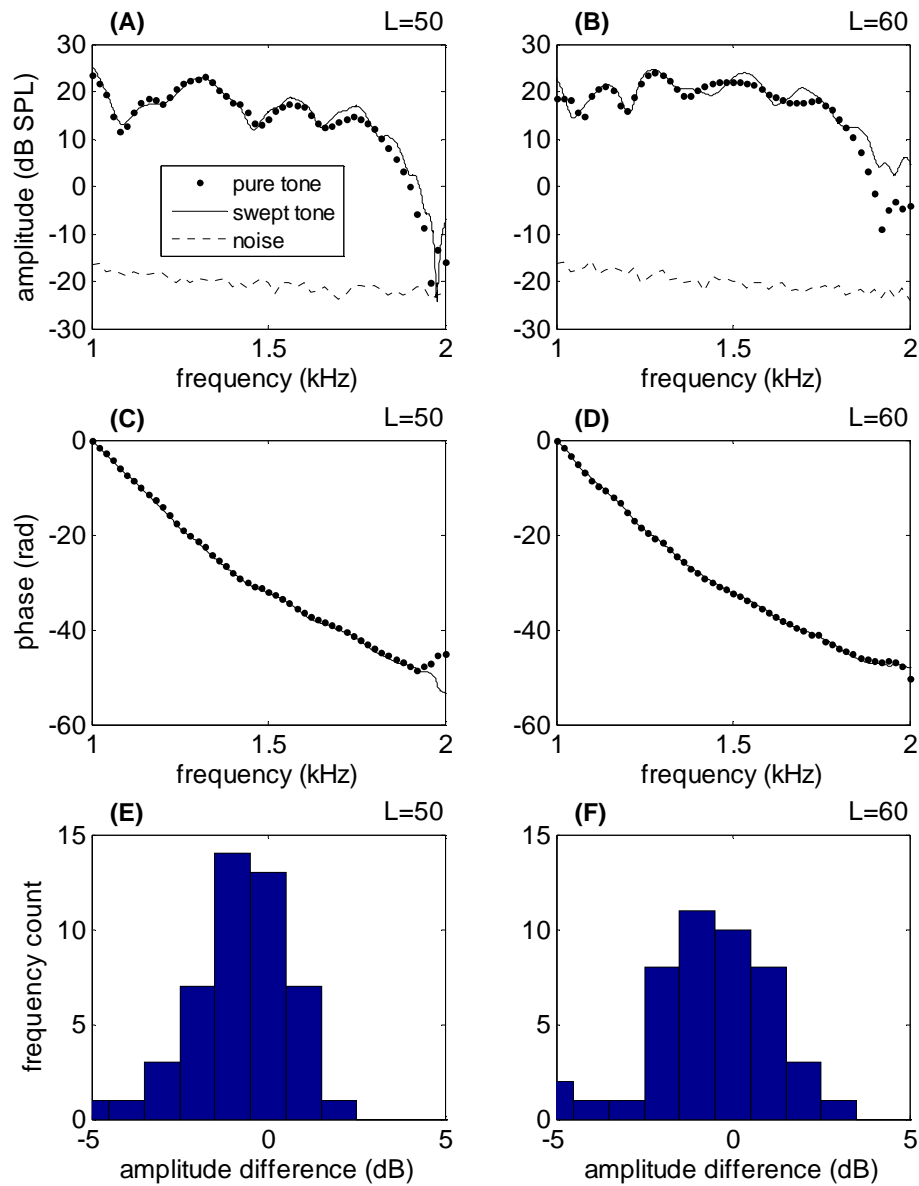
Figure 21. The amplitude spectra, phase spectra and the histograms of the
amplitude difference of the swept-tone and pure-tone SFOAEs under different
stimulus levels (subject # 1).

To investigate the consistency across subjects, the amplitude difference
(ΔA) and phase difference ($\Delta\theta$) averaged across frequency were shown in

Figure 22 (A)-(B) under two different stimulus levels (L). In panel (A), the averaged amplitude difference was less than 0.5 dB for both stimulus levels. A t-test was performed on ΔA and no significant L effects were found (p=0.47). A similar analysis was performed on the phase difference $\Delta \theta$ and the results showed that the phase difference was not significantly different from 0 for both stimulus levels. The findings indicated that the measurements with two different methods were identical across ears.

To further examine the similarity of the results, the absolute difference averaged across frequency ($|\Delta A|$) and correlation coefficient (r) of the amplitudes of pure-tone and swept-tone SFOAEs were compared in Figure 22 (C)-(D). It was observed that the $|\Delta A|$ across subjects had a mean of about 2 dB and a standard deviation of about 0.5 dB. The mean of the averaged r was around 0.9, and slight larger deviation was found for $L = 60$ dB SPL. No difference was found for $|\Delta A|$ and r between the 2 stimulus levels.
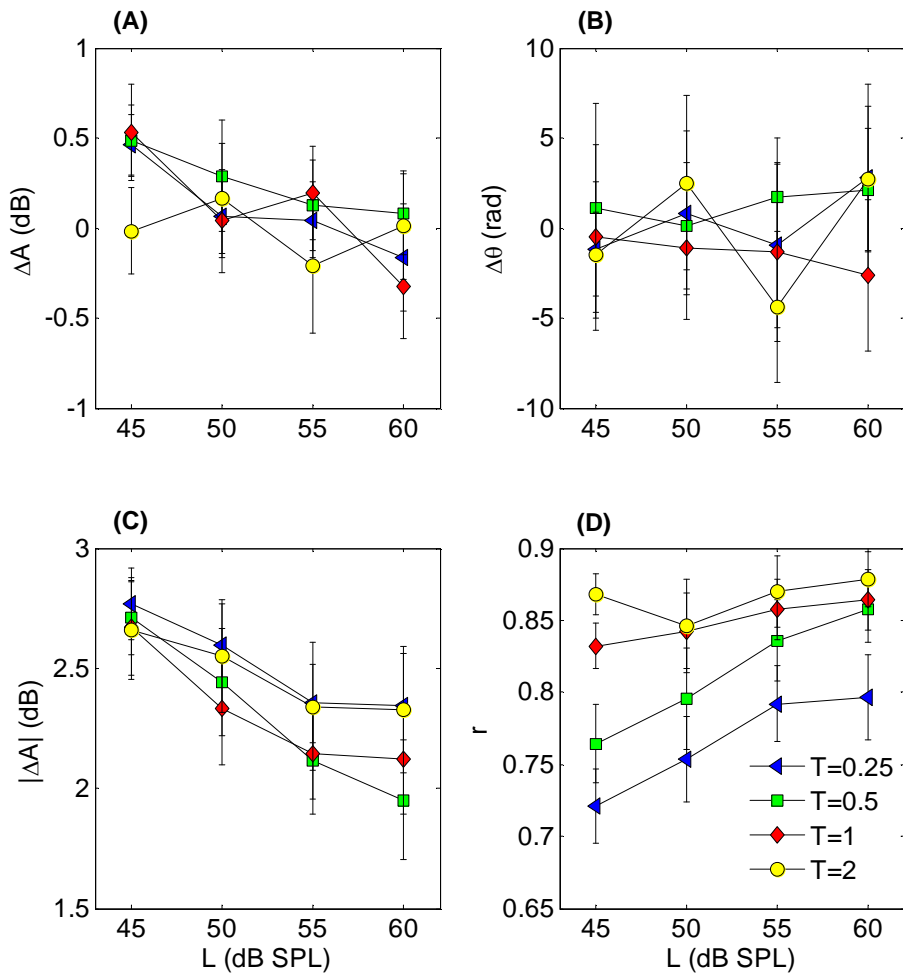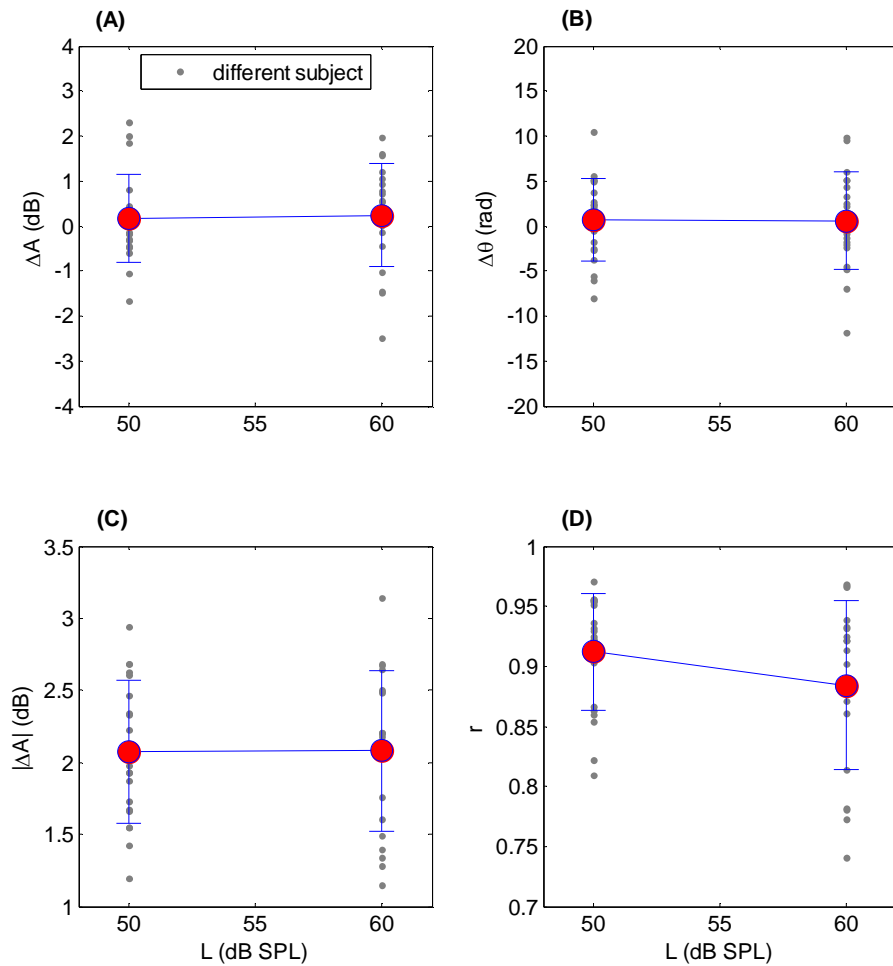
Figure 22. The amplitude difference (ΔA), phase difference ($\Delta\theta$), absolute amplitude difference ($|\Delta A|$) and correlation coefficient (r) between the swept-tone and pure-tone SFOAEs (n=20).

3.2.5 Further observations of swept-tone SFOAEs

A common observation of SFOAEs in normal-hearing subjects was the distinct fine structures that differed from subject to subject. The swept-tone SFOAE method allowed more detailed description of the fine structures at a frequency-resolution of less than 4 Hz across a wide frequency range. Analyzing

the features of the SFOAEs may provide new information for the understanding of cochlear mechanics.

A typical fine structure was bound by two neighboring notches (Figure 23) and the peak of the fine structure was obtained by finding the local maximum. The width of a particular fine structure (Δf) was determined by the frequency difference between the two neighboring notches ($\Delta f = f_{n2} - f_{n1}$), and the center frequency (f_c) was calculated by the mean $f_c = (f_{n1} + f_{n2})/2$. The peaks might not be always at the center of the fine structures. An eligible fine structure was qualified when the peak was 3 dB larger than either notch, and was confirmed by visual inspections.
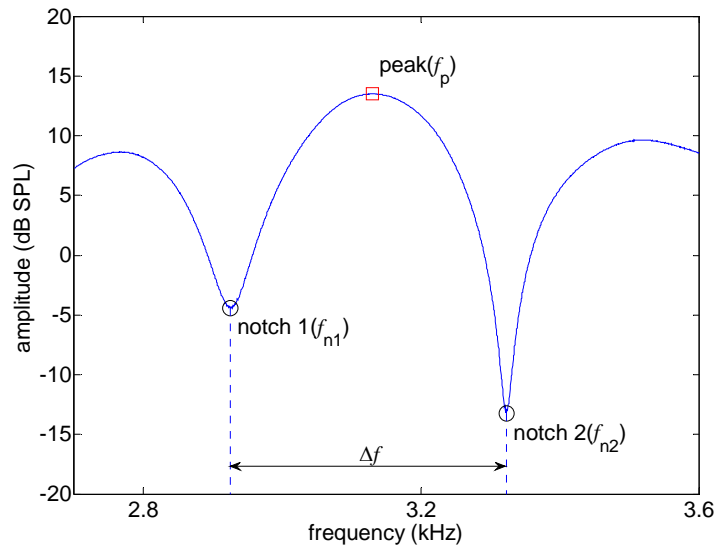
Figure 23. A typical fine structure of the swept-tone SFOAEs.

A) Width of fine structures

For each subject, fine structures were searched from 0.5 to 10 kHz in the amplitude spectrum of the swept-tone SFOAEs. The width of the fine structures (

75

Δf) as a function of the center frequency f_c from 20 subjects were plotted

together in Figure 24 for 4 different stimulus levels. A common trend observed

across level was that the Δf expanded as the f_c increased. The Δf was around 0.2

kHz when the center frequency was 1 kHz, and increased up to 1.2 kHz at10 kHz.

For each stimulus level, the scatter plot of the width-frequency function was fit

with a straight line, and the growing trend was shown in the last panel. The slopes

of the regression lines increased from 0.055 to 0.062 when the stimulus level was

from 45 to 60 dB SPL. A two-way ANOVA with f_c and L as IVs was performed

on the Δf . Significant effects of both IVs and interactions between them were

found ($p < 0.001$).
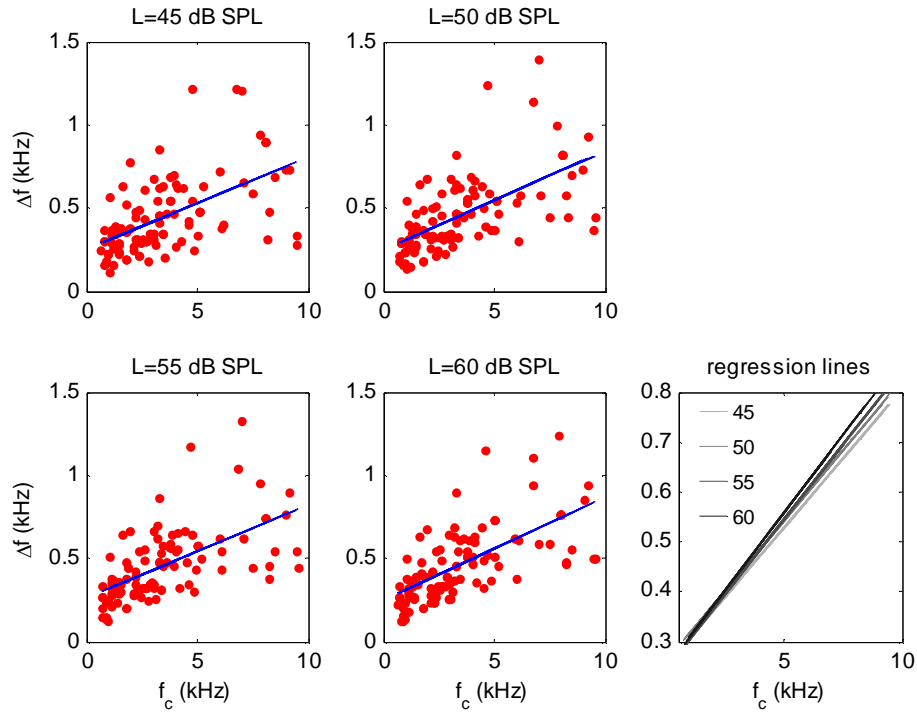
Figure 24. The width of the fine structure (Δf) as a function of frequency (f_c) under different stimulus levels (n=20).

B) Frequency shift of notches and peaks

In the above observation, the fine structure became wider as the stimulus level increased. In other words, there were fewer fine structures at higher stimulus levels and there ought to be frequency change of the remaining fine structures. The frequency shift of the peaks (or the notches) as a function of the stimulus level was therefore measured.

The frequency shifts of the notches and peaks from the lowest level (45 dB SPL) were calculated for all the subjects (Figure 25). A negative frequency change was found as the stimulus level increased. When the stimulus level increased from 45 t0 60 dB SPL, the total frequency shift was about 40 Hz for the

notches and it was about 50 Hz for the peaks. A one-way ANOVA was performed

(DV: the frequency shift; IV: the stimulus level) for the notches and peaks

respectively. P-values less than 0.001 (for both the notches and peaks) indicated

that the observed frequency shifts were statistically significant ($p < 0.001$).
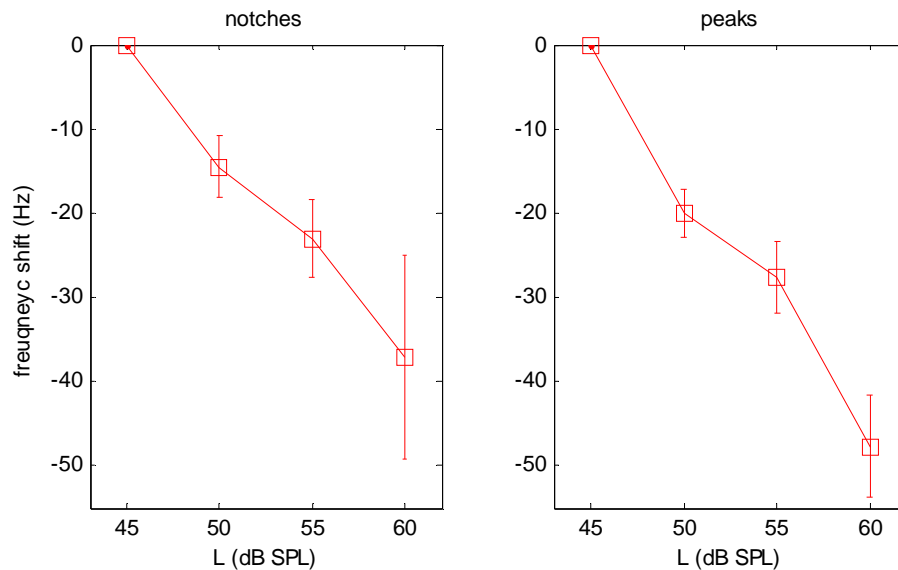
Figure 25. The frequency shifts of the notches and peaks as functions of stimulus

level (n=20).

3.3 Discussion

3.3.1 The swept tone method

This experiment has shown that swept tones could be used to evoke

SFOAEs. The use of swept tones allows an efficient, wideband and high-

definition measurement of SFOAEs. In this study, it took about 15-20 min to

measure SFOAEs from 1 to 2 kHz using pure tones with a frequency resolution of

50 points/octave, while it took only 1-2 minutes for the swept-tone method to

measure SFOAEs from 0.5 to 10 kHz with a much higher resolution using swept

tones. For pure-tone SFOAE measurements of other studies, it also took half an

78

hour or even longer to obtain enough number of frequencies within one or two

frequency octaves (Cheatham et al., 2011; Kalluri & Shera, 2007a; Schairer &

Keefe, 2005; Schairer et al., 2003). Although using clicks to measure TEOAEs

has the similar efficiency, the OAE frequency range is restricted to below 4 kHz

and the frequency resolution is also poorer than the swept-tone SFOAEs (Kemp,

1986).

The three-interval paradigm developed by Keefe (1998) was used to

generate the swept-tone SFOAEs in this study. The paradigm makes use of the

two-tone suppression phenomenon in cochlear mechanics to obtain a nonlinear

portion of the SFOAEs from a subtraction between the last interval with two-tone

stimuli and the first two intervals containing only one of the stimulus signals.

Compared with the suppression method used by other investigators (Kalluri &

Shera, 2007a; Kemp & Chum, 1980), the three-interval paradigm is more

beneficial in measuring swept-tone SFOAEs because it has advantage in

eliminating the impact of the stimulus artifacts. Note that when two tones are

presented simultaneously, they can produce mutual suppression on the SFOAEs

of each stimulus. Thus, the residual contains the nonlinear portion of both the s_1-

and s_2-evoked SFOAEs. The frequency-time functions of the two OAE

components are two parallel lines, similar to their evoking stimuli in Figure 13. A

conventional bandpass filter cannot separate the two emissions because they are

largely overlapped in frequency over the entire recording time. Therefore, a

tracking filter (Raja Kumar & Pal, 1990) whose center frequency adaptively

follows the instantaneous frequency the SFOAE evoked by the s_1 can effectively

remove the background noise and the SFOAE induce by the s_2. Since the level of

the suppressor s_2 was as high as 80 dB SPL and system distortions may be

involved at such a high level, the SFOAEs evoked by s_2 were not analyzed in this

study.

TEOAE frequencies are not strictly controlled by the stimulus. The

recording window is extended up to 20 ms after the stimulus, thus any vibrational

energy including multiple reflections simultaneous excitations from multiple

locations is collected (Kemp, 1978). Swept tone SFOAEs only contain the energy

generated by the cochlea immediately following the stimulus frequency, thus free

from multiple reflections and contribution from places other than the CF location.

Thus, swept tone SFOAEs may be more indicative of the cochlear functions

across the whole cochlea than the TEOAEs.

3.3.2 Features of SFOAEs

A) Global shape

The spectral pattern of the SFOAE amplitudes is unique to each subject.

The uniqueness is demonstrated in two aspects: global shape and fine structures.

The global shape refers to the change of the baseline amplitude across frequency.

Although the stimulus level has been equalized at the TM, the global shape of the

SFOAEs is not flat with broad peaks and valleys presented at various frequency

regions (Figure 16, Figure 17). Studies showed that the non-flat global shape of

the SFOAE amplitude could be due to the transfer function of the middle ear

(Paul Avan et al., 2000; Osterhammel et al., 1993) and the functional status of the

OHCs (Walsh & McGee, 1990). The global shape of the SFOAE amplitude of the

normal hearing subject shows peaks around 1 and 4 kHz (Figure 16 and 17). Such peaks can be observed in the middle-ear transfer function from other studies (Gan et al., 2007; Aibara et al., 2001). In addition to the contribution of the middle ear transfer function, the reduction of SFOAE amplitude in relatively broad frequency regions may be attributable to altered or even damaged OHC functions. Therefore, if the global SFOAE amplitude pattern can be normalized by the middle ear transfer characteristics, it could be useful for early identification of cochlear lesions.

B) Fine structures

One remarkable feature of the amplitude spectrum of the swept-tone SFOAE is the presence of the fine structures (or ripples), defined by repeatedly appearing peaks and notches across frequencies (Figure 16, Figure 17). The high frequency-resolution of the swept-tone SFOAE spectrum makes the analysis of fine structures more accurate and reliable. Fine structures of SFOAEs were observed in many other studies (Shera & Guinan, 2003; Talmadge et al., 2000; Zweig & Shera, 1995). These authors attribute the fine structures in OAEs to the existence of multiple reflections of traveling waves in the cochlea. SFOAEs are thought to be originated from coherent reflections within the peak of the travelling wave and the interference between the forward and reflected travelling waves could enhance or cancel the emissions depending on their phase relation (Schairer et al., 2006; Zweig & Shera, 1995). The coherent reflection is a linear scattering of the forward travelling wave due to the impedance irregularities within a localized region of the peak of the traveling wave (Shera & Guinan, 1999) and it

is the dominant source of SFOAE generation at low stimulus levels (Goodman et al., 2003; Zweig & Shera, 1995). Therefore, the fine structures become less pronounced as the stimulus level increases (Figure 16). It was also observed that the fine structures are narrow at low frequencies and wider at high frequencies (Figure 24). It has been reported that the ratio of the fine-structure width over the center frequency was roughly a constant ranging from 5% to 15% across frequency (Talmadge et al., 1998; Zweig & Shera, 1995). Such a constant ratio is also called the spectral regularity (Zweig & Shera, 1995). In the present study, the ratio, which is also the slope of the straight line in Figure 24, was about 6% which is consistent with other studies.

Studies showed that the fine structures of SFOAEs are correlated with the microstructures of the hearing threshold (Heitmann et al., 1996), and the presence of the fine structures is closely related to the normal activities of the OHCs (Henin et al., 2011; Rao & Long, 2011). Therefore, fine structures could be used as indicators of the normal OHC activities (Wagner, Plinkert, et al., 2008).

C) Phase-frequency functions

While the amplitude spectrum demonstrated considerate variation with the presence of fine structures, the phase-frequency function of the swept-tone SFOAEs shows a relatively simple pattern (Figure 16, Figure 17). The SFOAE phase showed a consistent decrease with increasing frequency. The slope of the phase-frequency curve is steeper for low frequencies, and shallower at high frequencies. These results are consistent with the findings of other studies (Bentsen et al., 2011; Shera & Guinan, 2003; Shera et al., 2002). The slopes of

82

phase-frequency function represent the group delays of the OAEs which can be interpreted as the round-trip travel time of the SFOAEs. Therefore, the group delay is longer for low frequencies and shorter for high frequencies. An explanation is that high-frequency SFOAEs are generated at the base of the basilar membrane and it takes shorter time for them to travel back to the ear canal; while low-frequency SFOAEs are generated at the apical sites and longer time is needed for them to travel back.

3.3.3 Effects of the stimulus level

As the stimulus level increased, the fine structures demonstrated a series of changes, such as the increase in the frequency spacing between neighboring notches (Figure 24) and the downward shift of the notch (or peak) frequencies (Figure 25). The increased of the width may be explained by the involvement of larger region of cochlear irregularities during the reflection at the peak of the travelling wave (Zweig & Shera, 1995). The downward shift of the notch (peak) frequency may be a result of the disappearance of fine structures in the low frequencies at high stimulus levels. Since the fine structures are narrower at low frequencies, some of the fine structures are most likely to disappear as they expand at high stimulus levels. As a result, the remaining fine structures show a shift towards low frequency.

A noteworthy finding of this study is that the phase-frequency function is level-dependent. As the stimulus level increases, the phase-frequency curves become progressively shallower (Figure 16). Such a level dependency agrees with the results of other studies (Choi et al., 2008; Shera & Guinan, 2003). When the

83

phase-frequency function is expressed as group delays, shallower phase functions

can be translated into shorter group delays at higher levels. The reduction in the

group delays may be explained by the shift of the travelling wave peak towards

the base at high stimulus levels (Recio et al., 1998; Ruggero et al., 1997). The

flattening of the phase function also correlates with the widening of the fine

structures.

3.3.4 Effects of the sweeping speed

Since the frequency was swept across the same range (from 0.5 to 10 kHz),

different durations correspond to different sweeping speeds. In our study, no

significant effects of the sweeping speed are found on either the amplitude or the

phase spectrum of the swept-tone SFOAEs (Figure 17). A possible explanation is

that the SFOAEs at different frequencies are generated independently, production

of SFOAE at one frequency does not temporally influence the SFOAEs

generation of the next frequency. The results show that the swept tone SFOAEs

are similar to pure tone evoked SFOAEs, thus for the fastest frequency sweeping

speed (100 kHz/0.25s), the cochlea can still follow the frequency change in the

stimulus and respond in a similar way as pure tones. Similarly, Kalluri and Shera

(2007b) varied the duration of the click stimulus and found no significant change

of the CEOAE amplitude. Further assumptions can be made based on the

insignificant effects of the sweeping speed: 1) since pure tones can be considered

as swept tones with a sweeping speed of zero, it could be assumed that SFOAEs

evoked by pure tones and swept tones would be comparable or equivalent. This

assumption was confirmed by the findings of this study; 3) since clicks can be

84

considered as swept tones with an infinite sweeping speed, the results of this study could possibly be consistent with TEOAEs measured under the same signal conditions. However, differences may be expected because the TEOAEs are largely produced after the cessation of the click thus containing various "free energy" from multiple reflections and simultaneous contribution from multiple locations. Whether the swept-tone SFOAEs are equivalent to TEOAEs needs to be tested in future research.

3.3.5 Test-retest reliability

Measuring the test-retest reliability of the swept-tone SFOAEs is a crucial step before the clinical applications, so that any changes in the results can be considered as evidence of cochlear functional change instead of measurement errors. There are many studies investigating the reliability of DPOAEs (Valero & Ratnam, 2011; Wagner, Heppelmann, et al., 2008; Parazzini, Galloni, et al., 2006; Parazzini, Wilson, et al., 2006; Zhao & Stephens, 1999) and TEOAEs (Keppler et al., 2010; Marshall & Heller, 1996; Franklin et al., 1992), but the reliability of SFOAEs are rarely explored. In this study, the reliability of the swept-tone SFOAEs was accessed under various signal conditions in different subjects. The results show that the amplitude and phase differences between the two repeated measures are not significantly different from 0 across all signal conditions and subjects (Figure 20). The mean absolute amplitude difference is less than 2.5 dB and the mean correlation coefficient was larger than 0.7 which is similar to the findings in the repeated measurements of DPOAEs (Sockalingam et al., 2007; Zhao & Stephens, 1999). Results from the present study show overall good test-

retest reliability over the entire frequency range (0.5-10 kHz), especially at high stimulus levels. Such level benefit can be explained by the higher signal to noise ratio at higher stimulus levels so that the SFOAEs are less likely to be affected by noises. Similar level effects were also found in other studies (Parazzini, Galloni, et al., 2006; Franklin et al., 1992). In Figure 20(D), the correlation coefficients between measurements are significantly lower for swept tones of shorter duration, especially for $T = 0.25$ and 0.5 s. This may be because the total time available for averaging was shorter for SFOAEs evoked by shorter swept tones during the experiment.

3.3.6 Consistency with pure-tone SFOAEs

It is also important to verify that the results of the swept tone SFOAEs are accurate. One way to verify the accuracy is to compare the results with the well-accepted method. In this study, the results of swept-tone SFOAEs are compared with the standard method using pure tones under the same signal condition. The comparison in this study reveals good agreement between the results of the two methods (Figure 21). The averaged differences (in both amplitude and phase) of the two methods are not significantly different from 0 across all the participants and signal levels (Figure 22). The absolute amplitude different is less than 2.5 dB with a correlation coefficient as high as 0.9. As discussed previously, the sweeping speed has no significant effects on the measurement of swept-tone SFOAEs and pure tones can be considered as swept tones with infinitely long durations. This result suggests that the SFOAEs obtained by pure tones are equivalent to SFOAEs measured by swept tones, especially with long durations.

86

Similarly, Long et al (2008) used continuously sweeping primaries (similar to swept tones) to measure DPOAEs and found that the results were nearly the same as the DPOAEs evoked by pure tones. They made their comparison over the frequency range of 0.5-2.5 kHz. The present study extended the comparison range up to 10 kHz. Kalluri and Shera (2001) compared the unmixed reflection component of DPOAE with the corresponding SFOAE and found that they were almost equivalent in both baseline amplitude and fines structures. Later (Kalluri & Shera, 2007b), they found that TEOAEs and SFOAEs were nearly identical at low and moderate stimulus intensities. These evidences suggest that, not only SFOAEs evoked by different types of evoking tones are identical, but also different types of OAEs may be comparable since they all originate from the OHC activities.

3.3.7 Clinical implications

Based on the results observed in this study, the swept-tone SFOAE method is a practical alternative to overcome the limitations of current OAE measurements. The use of swept tones allows OAEs to be efficiently measured in a wide frequency band and with a high frequency resolution. The measure of SFOAEs could help to provide a more precise mapping between the obtained OAE results and the functional status of cochlear OHCs, or so-called cochleogram. The good test-retest reliability ensures that the SFOAEs can be reproduced over time, so that any changes may be attributed to alterations of cochlear functions. The comparison with pure-tone SFOAEs indicates that the results of the swept-tone SFOAEs are accurate and consistent with standard methods. Therefore, the

swept-tone SFOAE method is recommended to supplement or even to replace the

current methods for routine OAE testing in the clinic.

Chapter 4

SUMMARY AND FUTURE RESEARCH

4.1 Summary

The purpose of this dissertation is to address two major concerns of current OAE measurements in the clinic: the inaccuracy in the signal calibration, and the limited bandwidth and efficiency of currently OAE methods. The approach was to use a swept tone stimuli to predict the sound pressure at the TM and to record SFOAEs efficiently.

In experiment I, an in-situ calibration was used to accurately control the stimulus levels that enter the middle ear. The experiment used a transmission line model to predict the eardrum pressure, so that the direct measurement of the TM pressure was not necessary. Then the benefits of the in-situ calibration on OAE measurements were explored. The results showed that the prediction of the TM pressure was successful in both the tube and in five human subjects. Compared with no calibration and probe calibration, the in-situ calibration was not affected by the standing-wave problem and it could obtain the most reliable results in OAE measurements. The clinical implication is that the in-situ calibration should be recommended during the OAE measurement to improve the reliability of the results.

In experiment II, the swept-tone SFOAE method was developed to solve the second problem. The experiment employed a three-interval paradigm and an adaptive filter to extract the swept-tone SFOAEs and the performance were tested under different signal conditions. The results showed that wideband SFOAEs

89

could be obtained with high frequency resolution within a short period of time. The swept-tone SFOAEs showed excellent test-retest reliability in repeated measurements with a 30 min interval. The results obtained with the swept tone approach were nearly equivalent with the standard method using pure tones under various signal conditions. The swept-tone method also allowed for more precise analysis of the features of SFOAEs, such as the fine structures and growth functions. The clinical implication is that the swept-tone SFOAEs may be an excellent alternative to current methods to provide reliable and detailed measurement of cochlear function over an extended frequency range.

4.2 Future research

One limitation of this study is that there were no hearing-impaired (HI) subjects recruited in the experiments. Since whether the normal-hearing (NH) group and HI group can be correctly identified is the main task of OAE screening, the future research can be extended to include the NI group. If two groups can be identified successfully, future research will explore whether swept tone SFOAE technique can precisely identify the frequency region of the hearing loss or even the degree of the hearing impairment.

SFOAE is defined as the OAEs evoke by a single tone. However, in practice, an additional suppressor tone has to be used to extract the emissions. With the techniques of the swept tone and the tracking filter, it is possible to obtain the SFOAEs without the second tone. This idea will be tried out in extended studies after this dissertation project.

REFERENCES

Aibara, R., Welsh, J. T., Puria, S., & Goode, R. L. (2001). Human middle-ear sound transfer function and cochlear input impedance. *Hear Res, 152*(1-2), 100-109.

Allen, J. B. (1986). Measurement of eardrum acoustic impedance. In J. H. JB Allen, A Hubbard, ST Neely, A Tubis (Ed.), *Peripery auditory mechanisms* (pp. 43-51). New York: Springer-Verlag.

Attias, J., Horovitz, G., El-Hatib, N., & Nageris, B. (2001). Detection and Clinical Diagnosis of Noise-Induced Hearing Loss by Otoacoustic Emissions. *Noise Health, 3*(12), 19-31.

Avan, P., Bonfils, P., Loth, D., & Wit, H. P. (1993). Temporal patterns of transient-evoked otoacoustic emissions in normal and impaired cochleae. *Hear Res, 70*(1), 109-120.

Avan, P., Büki, B., Maat, B., Dordain, M., & Wit, H. P. (2000). Middle ear influence on otoacoustic emissions. I: Noninvasive investigation of the human transmission apparatus and comparison with model results. *Hearing Research, 140*(1–2), 189-201.

Bentsen, T., Harte, J. M., & Dau, T. (2011). Human cochlear tuning estimates from stimulus-frequency otoacoustic emissions. *129*(6), 3797-3807.

Bian, L. (2004). Cochlear compression: effects of low-frequency biasing on quadratic distortion product otoacoustic emission. *J Acoust Soc Am, 116*(6), 3559-3571.

Bian, L., Chertoff, M. E., & Miller, E. (2002). Deriving a cochlear transducer function from low-frequency modulation of distortion product otoacoustic emissions. *J Acoust Soc Am, 112*(1), 198-210.

Bian, L., & Scherrer, N. M. (2007). Low-frequency modulation of distortion product otoacoustic emissions in humans. *J Acoust Soc Am, 122*(3), 1681.

Brownell, W., Bader, C., Bertrand, D., & de Ribaupierre, Y. (1985). Evoked mechanical responses of isolated cochlear outer hair cells. *Science, 227*(4683), 194-196.

Brummett, R. E. (1980). Drug-induced ototoxicity. *Drugs, 19*(6), 412-428.

Burke, S. R., Rogers, A. R., Neely, S. T., Kopun, J. G., Tan, H., & Gorga, M. P. (2010). Influence of calibration method on distortion-product otoacoustic emission measurements: I. test performance. *Ear Hear, 31*(4), 533-545.

Burns, E. M., Arehart, K. H., & Campbell, S. L. (1992). Prevalence of spontaneous otoacoustic emissions in neonates. *The Journal of the Acoustical Society of America, 91*(3), 1571-1575.

Cheatham, M. A., Katz, E. D., Charaziak, K., Dallos, P., & Siegel, J. H. (2011). Using Stimulus Frequency Emissions to Characterize Cochlear Function in Mice. *AIP Conference Proceedings, 1403*(1), 383-388.

Chen, S., & Bian, L. (2011). Cochlear Gain Control Estimated from Distortion Product Otoacoustic Emissions Evoked by Amplitude Modulated Tones. *AIP Conference Proceedings, 1403*(1), 601-607.

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*(5), 2651-2669.

Dalmont, J. P. (2001). Acoustic Impedance Measurement, Part II: a New Calibration Method. *Journal of Sound and Vibration, 243*(3), 441-459.

Davis, B., Qiu, W., & Hamernik, R. P. (2004). The use of distortion product otoacoustic emissions in the estimation of hearing and sensory cell loss in noise-damaged cochleas. *Hearing Research, 187*(1–2), 12-24.

Deltenre, P., Mansbach, A. L., Bozet, C., Christiaens, F., Barthelemy, P., Paulissen, D., & Renglet, T. (1999). Auditory neuropathy with preserved cochlear microphonics and secondary loss of otoacoustic emissions. *Audiology, 38*(4), 187-195.

Dhar, S., Talmadge, C. L., Long, G. R., & Tubis, A. (2002). Multiple internal reflections in the cochlea and their effect on DPOAE fine structure. *J Acoust Soc Am, 112*(6), 2882-2897.

Dirks, D. D., & Kincaid, G. E. (1987). Basic acoustic considerations of ear canal probe measurements. *Ear Hear, 8*(5 Suppl), 60S-67S.

Fobel, O., & Dau, T. (2004). Searching for the optimal stimulus eliciting auditory brainstem responses in humans. *J Acoust Soc Am, 116*(4 Pt 1), 2213-2222.

Franklin, D. J., McCoy, M. J., Martin, G. K., & Lonsbury-Martin, B. L. (1992). Test/retest reliability of distortion-product and transiently evoked otoacoustic emissions. *Ear Hear, 13*(6), 417-429.

Gan, R., Reeves, B., & Wang, X. (2007). Modeling of Sound Transmission from Ear Canal to Cochlea. *Annals of Biomedical Engineering, 35*(12), 2180-2195.

Gilman, S., & Dirks, D. (1985). Acoustics of ear canal measurement of eardrum SPL. *78*(S1), S12-S12.

Gilman, S., & Dirks, D. (1986). Acoustics of ear canal measurement of eardrum SPL in simulators. *J Acoust Soc Am, 80*(3), 783-793.

Goodman, S. S., Fitzpatrick, D. F., Ellison, J. C., Jesteadt, W., & Keefe, D. H. (2009). High-frequency click-evoked otoacoustic emissions and behavioral thresholds in humans. *J Acoust Soc Am, 125*(2), 1014-1032.

Goodman, S. S., Withnell, R. H., & Shera, C. A. (2003). The origin of SFOAE microstructure in the guinea pig. *Hear Res, 183*(1-2), 7-17.

Hamernik, R. P., Ahroon, W. A., Jock, B. M., & Bennett, J. A. (1998). Noise-induced threshold shift dynamics measured with distortion-product otoacoustic emissions and auditory evoked potentials in chinchillas with inner hair cell deficient cochleas. *Hearing Research, 118*(1-2), 73-82.

Hatzopoulos, S., Cheng, J., Grzanka, A., Morlet, T., & Martini, A. (2000). Optimization of TEOAE recording protocols: a linear protocol derived from parameters of a time-frequency analysis: a pilot study on neonatal subjects. *Scand Audiol, 29*(1), 21-27.

Heitmann, J., Waldmann, B., & Plinkert, P. K. (1996). Limitations in the use of distortion product otoacoustic emissions in objective audiometry as the result of fine structure. *Eur Arch Otorhinolaryngol, 253*(3), 167-171.

Henin, S., Thompson, S., Abdelrazeq, S., & 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*(4), 2068-2079.

Johnson, T. A., Neely, S. T., Kopun, J. G., Dierking, D. M., Tan, H., Converse, C., . . . Gorga, M. P. (2007). Distortion product otoacoustic emissions:

cochlear-source contributions and clinical test performance. *J Acoust Soc Am, 122*(6), 3539-3553.

Kalluri, R., & Shera, C. A. (2001). Distortion-product source unmixing: A test of the two-mechanism model for DPOAE generation. *109*(2), 622-637.

Kalluri, R., & Shera, C. A. (2007a). Comparing stimulus-frequency otoacoustic emissions measured by compression, suppression, and spectral smoothing. *J Acoust Soc Am, 122*(6), 3562-3575.

Kalluri, R., & Shera, C. A. (2007b). Near equivalence of human click-evoked and stimulus-frequency otoacoustic emissions. *121*(4), 2097-2110.

Keefe, D. H. (1998). Double-evoked otoacoustic emissions. I. Measurement theory and nonlinear coherence. *103*(6), 3489-3498.

Keefe, D. H., Ling, R., & Bulen, J. C. (1992). Method to measure acoustic impedance and reflection coefficient. *J Acoust Soc Am, 91*(1), 470-485.

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *J Acoust Soc Am, 64*(5), 1386-1391.

Kemp, D. T. (1986). Otoacoustic emissions, travelling waves and cochlear mechanisms. *Hear Res, 22*, 95-104.

Kemp, D. T. (2002). Otoacoustic emissions, their origin in cochlear function, and use. *Br Med Bull, 63*, 223-241.

Kemp, D. T., & Chum, R. A. (1980). Observation on the generation of stimulus frequency acoustic emissions – Two tone suppression. In G. van den Brink & F. A. Bilsen (Eds.), *Psychophysical, Physiological and Behavioural Studies in Hearing*. Delft: Delft Univ. Press.

Kemp, D. T., Ryan, S., & Bray, P. (1990). A guide to the effective use of otoacoustic emissions. *Ear Hear, 11*(2), 93-105.

Keppler, H., Dhooge, I., Maes, L., D'Haenens, W., Bockstael, A., Philips, B., . . . Vinck, B. (2010). Transient-evoked and distortion product otoacoustic emissions: A short-term test-retest reliability study. *Int J Audiol, 49*(2), 99-109.

Kossl, M., & Vater, M. (2000). Consequences of outer hair cell damage for otoacoustic emissions and audio-vocal feedback in the mustached bat. *J Assoc Res Otolaryngol, 1*(4), 300-314.

Kruger, B., & Ruben, R. J. (1987). The acoustic properties of the infant ear. A preliminary report. *Acta Otolaryngol, 103*(5-6), 578-585.

Kuronen, P., Sorri, M. J., Paakkonen, R., & Muhli, A. (2003). Temporary threshold shift in military pilots measured using conventional and extended high-frequency audiometry after one flight. *Int J Audiol, 42*(1), 29-33.

Levy, Y., & Azhari, H. (2007). Velocity measurements using a single transmitted linear frequency-modulated chirp. *Ultrasound Med Biol, 33*(5), 768-773.

Lewis, J. D., McCreery, R. W., Neely, S. T., & Stelmachowicz, P. G. (2009). Comparison of in-situ calibration methods for quantifying input to the middle ear. *J Acoust Soc Am, 126*(6), 3114-3124.

Long, G. R., Talmadge, C. L., & Lee, J. (2008). Measuring distortion product otoacoustic emissions using continuously sweeping primaries. *124*(3), 1613-1626.

Lonsbury-Martin, B. L., Harris, F. P., Stagner, B. B., Hawkins, M. D., & Martin, G. K. (1990). Distortion product emissions in humans. I. Basic properties in normally hearing subjects. *Ann Otol Rhinol Laryngol Suppl, 147*, 3-14.

Lonsbury-Martin, B. L., Martin, G. K., Probst, R., & Coats, A. C. (1988). Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy. *Hear Res, 33*(1), 69-93.

Lutman, M. E., & Deeks, J. (1999). Correspondence amongst microstructure patterns observed in otoacoustic emissions and Bekesy audiometry. *Audiology, 38*(5), 263-266.

Marshall, L., & Heller, L. M. (1996). Reliability of transient-evoked otoacoustic emissions. *Ear Hear, 17*(3), 237-254.

Marshall, L., Heller, L. M., & Westhusin, L. J. (1997). Effect of negative middle-ear pressure on transient-evoked otoacoustic emissions. *Ear Hear, 18*(3), 218-226.

Martin, G. K., Jassir, D., Stagner, B. B., & Lonsbury-Martin, B. L. (1998). Effects of loop diuretics on the suppression tuning of distortion-product otoacoustic emissions in rabbits. *104*(2), 972-983.

Martin, G. K., Stagner, B. B., & Lonsbury-Martin, B. L. (2010). Evidence for basal distortion-product otoacoustic emission components. *J Acoust Soc Am, 127*(5), 2955-2972.

McCreery, R. W., Pittman, A., Lewis, J., Neely, S. T., & 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*(1), 15-24.

Meenderink, S. W., & van der Heijden, M. (2011). Distortion product otoacoustic emissions evoked by tone complexes. *J Assoc Res Otolaryngol, 12*(1), 29-44.

Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A Model for the Prediction of Thresholds, Loudness, and Partial Loudness. *Journal of the Audio Engineering Society, 45*(4), 224-240.

Müller, S. M., Paulo. (2001). Transfer-Function Measurement with Sweeps. *J. Audio Eng. Soc, 49*(6), 443-471.

Neely, S. T., & Gorga, M. P. (1998). Comparison between intensity and pressure as measures of sound level in the ear canal. *J Acoust Soc Am, 104*(5), 2925-2934.

Neumann, J., Uppenkamp, S., & Kollmeier, B. (1994). Chirp evoked otoacoustic emissions. *Hear Res, 79*(1-2), 17-25.

Nozza, R. J., Sabo, D. L., & Mandel, E. M. (1997). A role for otoacoustic emissions in screening for hearing impairment and middle ear disorders in school-age children. *Ear Hear, 18*(3), 227-239.

Ohlemiller, K. K., & Siegel, J. H. (1992). The effects of moderate cooling on gross cochlear potentials in the gerbil: Basal and apical differences. *Hearing Research, 63*(1-2), 79-89.

Osterhammel, P. A., Nielsen, L. H., & Rasmussen, A. N. (1993). Distortion Product Otoacoustic Emissions The Influence of the Middle Ear Transmission. *Scand Audiol, 22*(2), 111-116.

Parazzini, M., Galloni, P., Brazzale, A. R., Tognola, G., Marino, C., & Ravazzani, P. (2006). Quantitative indices for the assessment of the repeatability of distortion product otoacoustic emissions in laboratory animals. *Biomedical Engineering, IEEE Transactions on, 53*(8), 1550-1556.

Parazzini, M., Wilson, H. K., Bell, S., Tognola, G., Ravazzani, P., & Lutman, M. E. (2006, Aug. 30 2006-Sept. 3 2006). *The Generation Mechanisms and Repeatability of 2F1 - F2 Distortion Product Otoacoustic Emissions: study on normally hearing subjects.* Paper presented at the Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE.

Patuzzi, R., & Moleirinho, A. (1998). Automatic monitoring of mechano-electrical transduction in the guinea pig cochlea. *Hear Res, 125*(1-2), 1-16.

Penner, M. J., & Zhang, T. (1997). Prevalence of spontaneous otoacoustic emissions in adults revisited. *Hear Res, 103*(1-2), 28-34.

Plinkert, P. K., Hemmert, W., Wagner, W., Just, K., & Zenner, H. P. (1999). Monitoring noise susceptibility: sensitivity of otoacoustic emissions and subjective audiometry. *Br J Audiol, 33*(6), 367-382.

Probst, R., Lonsbury-Martin, B. L., & Martin, G. K. (1991). A review of otoacoustic emissions. *89*(5), 2027-2067.

Puria, S. (2003). Measurements of human middle ear forward and reverse acoustics: Implications for otoacoustic emissions. *The Journal of the Acoustical Society of America, 113*(5), 2773-2789.

Raja Kumar, R. V., & Pal, R. N. (1990). Tracking of bandpass signals using center-frequency adaptive filters. *Acoustics, Speech and Signal Processing, IEEE Transactions on, 38*(10), 1710-1721.

Rao, A., & Long, G. R. (2011). Effects of aspirin on distortion product fine structure: interpreted by the two-source model for distortion product otoacoustic emissions generation. *J Acoust Soc Am, 129*(2), 792-800.

Reavis, K. M., Phillips, D. S., Fausti, S. A., Gordon, J. S., Helt, W. J., Wilmington, D., . . . Konrad-Martin, D. (2008). Factors affecting sensitivity of distortion-product otoacoustic emissions to ototoxic hearing loss. *Ear Hear, 29*(6), 875-893.

Recio, A., Rich, N. C., Narayan, S. S., & Ruggero, M. A. (1998). Basilar-membrane responses to clicks at the base of the chinchilla cochlea. *J Acoust Soc Am, 103*(4), 1972-1989.

Rogers, A. R., Burke, S. R., Kopun, J. G., Tan, H., Neely, S. T., & Gorga, M. P. (2010). Influence of calibration method on distortion-product otoacoustic emission measurements: II. threshold prediction. *Ear Hear, 31*(4), 546-554.

Roos, P. A., Reibel, R. R., Berg, T., Kaylor, B., Barber, Z. W., & Babbitt, W. R. (2009). Ultrabroadband optical chirp linearization for precision metrology applications. *Opt Lett, 34*(23), 3692-3694.

Ross, S. (1968). Impedance at the Eardrum, Middle-Ear Transmission, and Equal Loudness. *The Journal of the Acoustical Society of America, 43*(3), 491-505.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., & Robles, L. (1997). Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J Acoust Soc Am, 101*(4), 2151-2163.

Schairer, K. S., Ellison, J. C., Fitzpatrick, D., & Keefe, D. H. (2006). Use of stimulus-frequency otoacoustic emission latency and level to investigate cochlear mechanics in human ears. *120*(2), 901-914.

Schairer, K. S., Fitzpatrick, D., & Keefe, D. H. (2003). Input-output functions for stimulus-frequency otoacoustic emissions in normal-hearing adult ears. *J Acoust Soc Am, 114*(2), 944-966.

Schairer, K. S., & Keefe, D. H. (2005). Simultaneous recording of stimulus-frequency and distortion-product otoacoustic emission input--output functions in human ears. *The Journal of the Acoustical Society of America, 117*(2), 818-832.

Scheperle, R. A., Neely, S. T., Kopun, J. G., & Gorga, M. P. (2008). Influence of in situ, sound-level calibration on distortion-product otoacoustic emission variability. *J Acoust Soc Am, 124*(1), 288-300.

Schweitzer, V. G., Hawkins, J. E., Lilly, D. J., Litterst, C. J., Abrams, G., Davis, J. A., & Christy, M. (1984). Ototoxic and nephrotoxic effects of combined treatment with cis-diamminedichloroplatinum and kanamycin in the guinea pig. *Otolaryngol Head Neck Surg, 92*(1), 38-49.

Seixas, N. S., Kujawa, S. G., Norton, S., Sheppard, L., Neitzel, R., & Slee, A. (2004). Predictors of hearing threshold levels and distortion product otoacoustic emissions among noise exposed young adults. *Occup Environ Med, 61*(11), 899-907.

Shaffer, L. A., Withnell, R. H., Dhar, S., Lilly, D. J., Goodman, S. S., & Harmon, K. M. (2003). Sources and mechanisms of DPOAE generation: implications for the prediction of auditory sensitivity. *Ear Hear, 24*(5), 367-379.

Shera, C. A., & Guinan, J. J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *J Acoust Soc Am, 105*(2 Pt 1), 782-798.

Shera, C. A., & Guinan, J. J. (2003). Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning. *113*(5), 2762-2772.

Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc Natl Acad Sci U S A, 99*(5), 3318-3323.

Siegel, J. H. (1994). Ear-canal standing waves and high-frequency sound calibration using otoacoustic emission probes. *95*(5), 2589-2597.

Siegel, J. H., Cerka, A. J., Recio-Spinoso, A., Temchin, A. N., van Dijk, P., & 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*(4), 2434-2443.

Siegel, J. H., & Hirohata, E. T. (1994). Sound calibration and distortion product otoacoustic emissions at high frequencies. *Hear Res, 80*(2), 146-152.

Sockalingam, R., Lee Choi, J., Choi, D., & Kei, J. (2007). Test-retest reliability of distortion-product otoacoustic emissions in children with normal hearing: a preliminary study. *Int J Audiol, 46*(7), 351-354.

Starr, A., Sininger, Y., Nguyen, T., Michalewski, H. J., Oba, S., & Abdala, C. (2001). Cochlear receptor (microphonic and summating potentials, otoacoustic emissions) and auditory pathway (auditory brain stem potentials) activity in auditory neuropathy. *Ear Hear, 22*(2), 91-99.

Stavroulaki, P., Vossinakis, I. C., Dinopoulou, D., Doudounakis, S., Adamopoulos, G., & Apostolopoulos, N. (2002). Otoacoustic emissions for monitoring aminoglycoside-induced ototoxicity in children with cystic fibrosis. *Arch Otolaryngol Head Neck Surg, 128*(2), 150-155.

Stevens, K. N., Berkovitz, R., Gerald Kidd, J., & Green, D. M. (1987). Calibration of ear canals for audiometry at high frequencies. *The Journal of the Acoustical Society of America, 81*(2), 470-484.

Strube, H. W. (1989). Evoked otoacoustic emissions as cochlear Bragg reflections. *Hear Res, 38*(1-2), 35-45.

Subramaniam, M., Salvi, R. J., Spongr, V. P., Henderson, D., & Powers, N. L. (1994). Changes in distortion product otoacoustic emissions and outer hair cells following interrupted noise exposures. *Hearing Research, 74*(1-2), 204-216.

Suzuki, Y., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *J Acoust Soc Am, 116*(2), 918-933.

Talmadge, C. L., Tubis, A., Long, G. R., & Piskorski, P. (1998). Modeling otoacoustic emission and hearing threshold fine structures. *J Acoust Soc Am, 104*(3 Pt 1), 1517-1543.

Talmadge, C. L., Tubis, A., Long, G. R., & Tong, C. (2000). Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure. *108*(6), 2911-2932.

Telischi, F. F., Widick, M. P., Lonsbury-Martin, B. L., & McCoy, M. J. (1995). Monitoring cochlear function intraoperatively using distortion product otoacoustic emissions. *Am J Otol, 16*(5), 597-608.

Tognola, G., Grandori, F., & Ravazzani, P. (1997). Time-frequency distributions of click-evoked otoacoustic emissions. *Hear Res, 106*(1-2), 112-122.

Valero, M. D., & Ratnam, R. (2011). Reliability of distortion-product otoacoustic emissions in the common marmoset (Callithrix jacchus). *Hear Res, 282*(1-2), 265-271.

van Huffelen, W. M., Mateijsen, N. J. M., & Wit, H. P. (1998). Classification of Patients with Ménière's Disease Using Otoacoustic Emissions. *Audiology and Neurotology, 3*(6), 419-430.

Wagner, W., Heppelmann, G., Vonthein, R., & Zenner, H. P. (2008). Test-retest repeatability of distortion product otoacoustic emissions. *Ear Hear, 29*(3), 378-391.

Wagner, W., & Plinkert, P. K. (1999). The relationship between auditory threshold and evoked otoacoustic emissions. *Eur Arch Otorhinolaryngol, 256*(4), 177-188.

Wagner, W., Plinkert, P. K., Vonthein, R., & 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*(10), 1165-1172.

Walsh, E. J., & McGee, J. (1990). Frequency selectivity in the auditory periphery: similarities between damaged and developing ears. *Am J Otolaryngol, 11*(1), 23-32.

Weiss, T. F., & Leong, R. (1985). A model for signal transmission in an ear having hair cells with free-standing stereocilia. IV. Mechanoelectric transduction stage. *Hear Res, 20*(2), 175-195.

Whitehead, M. L., McCoy, M. J., Lonsbury-Martin, B. L., & Martin, G. K. (1995). Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. I. Effects of decreasing L2 below L1. *J Acoust Soc Am, 97*(4), 2346-2358.

Whitehead, M. L., Stagner, B. B., Lonsbury-Martin, B. L., & Martin, G. K. (1995). Effects of ear-canal standing waves on measurements of distortion-product otoacoustic emissions. *J Acoust Soc Am, 98*(6), 3200-3214.

Wilson, H. K., & Lutman, M. E. (2006). Mechanisms of generation of the 2f2-f1 distortion product otoacoustic emission in humans. *J Acoust Soc Am, 120*(4), 2108-2115.

Withnell, R. H., Jeng, P. S., Waldvogel, K., Morgenstein, K., & Allen, J. B. (2009). An in situ calibration for hearing thresholds. *J Acoust Soc Am, 125*(3), 1605-1611.

Withnell, R. H., & Yates, G. K. (1998). Enhancement of the transient-evoked otoacoustic emission produced by the addition of a pure tone in the guinea pig. *J Acoust Soc Am, 104*(1), 344-349.

Zhao, F., & Stephens, D. (1999). Test-retest variability of distortion-product otoacoustic emissions in human ears with normal hearing. *Scand Audiol, 28*(3), 171-178.

Zweig, G., & Shera, C. A. (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. *J Acoust Soc Am, 98*(4), 2018-2047.

APPENDIX A

CONSTRUCTION OF SWEPT TONES

The swept tone was constructed in the frequency domain and the construction steps included (Müller, 2001): (1) prepare the required amplitude spectrum (amplitude as a function of frequency); (2) prepare a phase spectrum according to the "frequency-time function" of the swept tone; (3) transfer the prepared amplitude and phase spectra by an inverse fast Fourier transform (FFT) to get the time waveform of the swept tone. For the first step, the amplitude spectrum can be prepared by constructing arbitrary amplitude-frequency functions. For the second step, the phase spectrum can be obtained by using the concept of "group delay", defined as "exactly at which time momentary (τ) a certain frequency (f) occurs". If the frequency is swept linearly from f_1 to f_2 when the time changes from 0 to T (in second), the τ-f relation is shown in Figure 26.
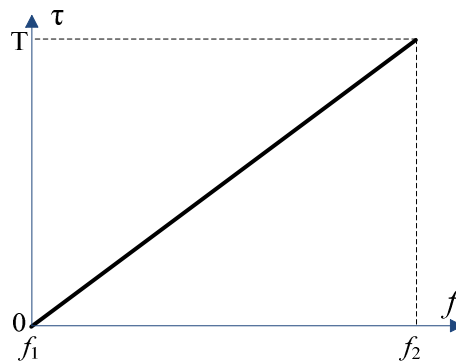

Figure 26. The relation between the frequency (f) and group delay (τ) in a linear swept tone.

The mathematical relation of Figure 26 could be expressed as $\tau = T(f - f_1)/(f_2 - f_1)$. On the other hand, group delay is generally defined as the rate of change of the total phase shift with respect to frequency:

$\tau = -d\varphi/(2\pi df)$. Therefore, the phase function can be obtained by taking the integral and finally we obtained:

$$\varphi(f) = -\frac{\pi T(f - f_1)^2}{f_2 - f_1} \tag{6}$$

With the prepared amplitude and phase spectra, the temporal waveform of the swept tone could be obtained by taking the inverse FFT. The resulted swept tone sweeps the frequency linearly from f_1 to f_2 within the duration of T. The advantage of the swept tone is that the amplitude of each individual frequency can be precisely controlled at the desired level. To avoid the switching noises caused by non-zero start and end of the stimulus, a half-cosine window with a time length of 20 ms was applied to the very beginning and end of the swept tone.

APPENDIX B

SOURCE CALIBRATION

A multiple-cavity method (Allen, 1986) was used for the source calibration in this study. In Allen's method, when the earphone was connected to an acoustic load (such as a tube or the ear canal), a Thevenin-equivalent circuit of the sound delivery system was shown in Figure 27 (P_s: source pressure; Z_s: source impedance; Z_L: load impedance; P_L: pressure measured at the load). According to Ohm's law, the pressures and impedances were related by the following equation:

$$\frac{P_s}{Z_s + Z_L} = \frac{P_L}{Z_L} \tag{7}$$
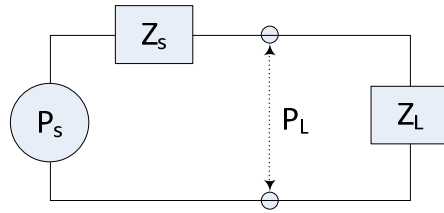
i.e.,

$$Z_L P_s - P_L Z_s = P_L Z_L \tag{8}$$

Figure 27. The Thevenin-equivalent circuit of the sound delivery system.

For a uniform tube with a length of L, the theoretical acoustic impedance (Z_L) can be calculated as (Keefe et al., 1992):

$$Z_L = -j\rho c \cot(2\pi fL / c) \tag{9}$$

where $j = \sqrt{-1}$, ρ is the air density, c is the sound speed in the air, f is frequency and L is the length of the tube. It could be observed in equation (9) that L is the only variable to determine Z_L (other parameters, such as ρ and c, are constants).

The tube length L could be either measured directly (from the earphone to the other closed end of the tube when the earphone was connected to the tube), or estimated from the probe pressure P_L. In the amplitude spectrum of P_L, a notch would be present at a certain frequency (f_0, dependent on L) because of the standing wave inside the tube. If the f_0 was measured in the spectrum of PL, the tube length L could be calculated as: $L = c / 4 f_0$.

In order to solve the two unknown variables P_s and Z_s in equation (8), the impedance Z_L in Figure 27 could be replaced by 5 uniform tubes with known impedances Z_i ($i = 1, 2, 3, 4, 5$) successively and the probe pressure P_i in response to the same wide-band swept-tone stimulus were measured. The results of each uniform tube could derivate an equation (8), and all the five equations together could be used to solve for the two unknown variables (P_s and Z_s) via a least mean squares method (Generally, two equations were enough to solve for two variables; using five equations in this study was to improve the accuracy of the solution):

$$\begin{cases} Z_1 P_s - P_1 Z_s = P_1 Z_1 \\ Z_2 P_s - P_2 Z_s = P_2 Z_2 \\ Z_3 P_s - P_3 Z_s = P_3 Z_3 \\ Z_4 P_s - P_4 Z_s = P_4 Z_4 \\ Z_5 P_s - P_5 Z_s = P_5 Z_5 \end{cases} \tag{10}$$

Where the pressure P_i ($i = 1, 2, 3, 4, 5$) was measured, and the impedance Z_i could be determined by equation (9). P_i and Z_i were both expressed as complex numbers as functions of frequency (expressed in the frequency domain).

Equation (10) was solved frequency by frequency. Consequently, the solutions of P_s and Z_s were also complex numbers as functions of frequency.

Instead of using five independent brass tubes in Allen and others' studies (Burke et al., 2010; Lewis et al., 2009; Scheperle et al., 2008), a plastic uniform tube with a moving end was used in this study (Figure 28). The moving end was achieved by using a piston that could move freely inside the tube. The advantage of the moving-end tube is that different tube lengths (*L*) could be obtained by moving the piston to different positions, without the requirement of moving the earphone to different tubes (which could introduce possible errors).
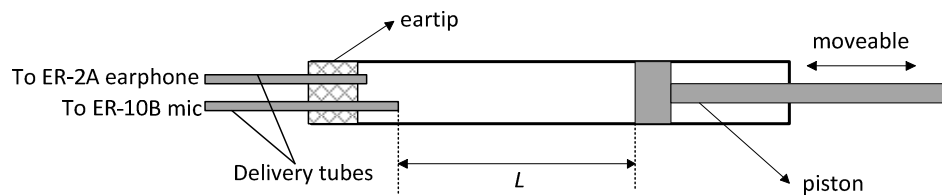

Figure 28. A plastic tube with a movable piston inside was used to calculate the source parameters.

APPENDIX C

OBTAINING THE EARDRUM PRESSURE

The two components (incident pressure P_+ and reflected pressure P_-) at

the entry of the ear canal could be rewritten as the polar form:

$$P_\pm = A_\pm(f) \angle \theta_\pm(f) \tag{11}$$

Where $A_\pm(f)$ and $\theta_\pm(f)$ were the amplitudes and phase (as functions of

frequency) of P_+ and P_-, respectively. The incident pressure P_+ travelled forward

along the ear canal and was measured as P_+' when it reached the eardrum (Figure

29). It could be assumed that P_+ and P_+' had the same amplitude (assuming no

energy absorption by the wall of the ear canal). However, their phase were

different since P_+' travelled a distance of L (the length of the ear canal) to arrive at

the eardrum. The phase difference $[\Delta\theta(f)]$ between P_+ and P_+' was frequency

dependent and related to the length of the ear canal L by $\Delta\theta(f) = 2\pi f L / c$ (c:

sound speed). The length of the ear canal L could be calculated from the

frequency of the first notch of the probe pressure $P_L(f_0)$ by $L = c / 4f_0$. Finally,

the phase of P_+' could be obtained by adding the phase shift $\Delta\theta(f)$ to the phase

of P_+. Similarly, the reflected pressure at the eardrum (P_-') could be obtained by

subtracting the same phase shift $\Delta\theta(f)$ from P_-. The corresponding equation

was:

$$P_\pm' = A_\pm(f) \angle [\theta_\pm(f) \pm \Delta\theta(f)] \tag{12}$$

Since the eardrum pressure was the sum of the incident and reflected

pressures at the eardrum position, the predicted eardrum pressure ($\hat{P}_E$) could be

obtained by:
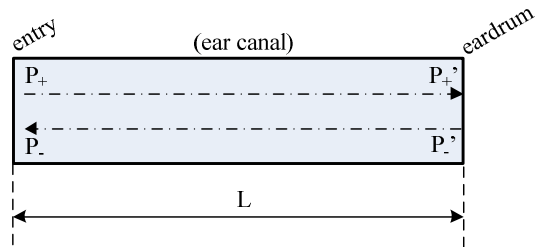
$$\hat{P}_E = P_+^{'} + P_-^{'} \tag{13}$$

Figure 29. The original and reflected pressures at the entry of the ear canal and at

the eardrum position.

APPENDIX D

CALIBRATION OF THE MICROPHONES

There were two microphones used in Experiment 1: the probe microphone and the eardrum microphone. The probe microphone was used to predict the eardrum pressure, while the eardrum microphone was used to measure the actual eardrum pressure. The predicted and measure the pressures would be compared to verify the reliability of the prediction.

The eardrum microphone must be calibrated in reference to the probe microphone before the two pressures could be compared. The main idea of the calibration was to let the two microphone measure the same pressure and the differences in their results were recorded (Figure 30). The same ER-7 tube used in the ear canal verification was attached to the eardrum microphone. The free end of the tube was placed close to the opening of the probe microphone so that the two microphones were measuring the same pressure. Then a swept tone stimulus (f_1=500 Hz, f_2=10 kHz, T=1 s and driving voltage=1 volt) was played by the earphone probe.

Figure 30. The diagram to calibrate the eardrum microphone with respect to the probe microphone.

If the response of the probe microphone was P_1 and it was P_2 for the eardrum microphone (P_1 and P_2 were expressed in the frequency domain), the ratio of the two responses was:

$$R = \frac{P_1}{P_2} \qquad (14)$$

If the reading of the eardrum microphone in response to any other pressure was P, then

the calibrated result (P') that was comparable to the probe microphone was:

$$P' = RP = \frac{P_1}{P_2}P \qquad (15)$$

APPENDIX E

THE TRACKING FILTER

There were two parts of swept-tone SFOAEs in the residual of the three-interval paradigm: one evoked by s_1 (s_1-OAE) and the other evoked by s_2 (s_2-OAE). The frequencies of the two OAEs were different and they changed over time in the same way as their evoking swept-tone stimulus [Figure 31(A)]. In this study, only the OAE by s_1 was analyzed, using a tracking-filter technique.

A tracking filter is narrow band-pass filter whose center frequency can change with time. The frequency of s_1-OAE, which could be estimated from the stimulus of s_1, was used as the center frequency of the tracking filter so that it could let go only the s_1-OAE and attenuate all other noises. To achieve this, the pole of the tracking filter (p_1) was placed at the frequency of s_1-OAE [Figure 31(A)]. Meanwhile, a zero (z_1) was placed at the frequency of s_2-OAE so that this part of OAE was eliminated and would not cause interference to the analysis of s_1-OAE. The pole and zero on the z-plane were shown in Figure 31(B). Consequently, the transfer function of the tracking filter is:

$$H(z) = G \frac{1 - z_1 z^{-1}}{1 - p_1 z^{-1}} \tag{16}$$

$$p_1 = r_p \exp(j\omega_1) \tag{17}$$

$$z_1 = r_z \exp(j\omega_2) \tag{18}$$

where ω_1 is the angular frequency of s_1-OAE, ω_2 is the angular frequency of s_2-OAE, r_p is the radius of the pole p_1, r_z is the radius of the zero z_1 and G is the gain. In this study, r_z was chosen to be 1 so that the tracking filter could eliminate the s_2-OAE to the

largest extent. The pole radius r_p was close to 1, dependent on the bandwidth (Δf) of the

tracking filter by:

$$r_p = 1 - \frac{\Delta f}{f_s}\pi \tag{19}$$

Where f_s is the sampling frequency. The gain G was determined by the following

equation to ensure unity gain at the s_1-OAE frequency:

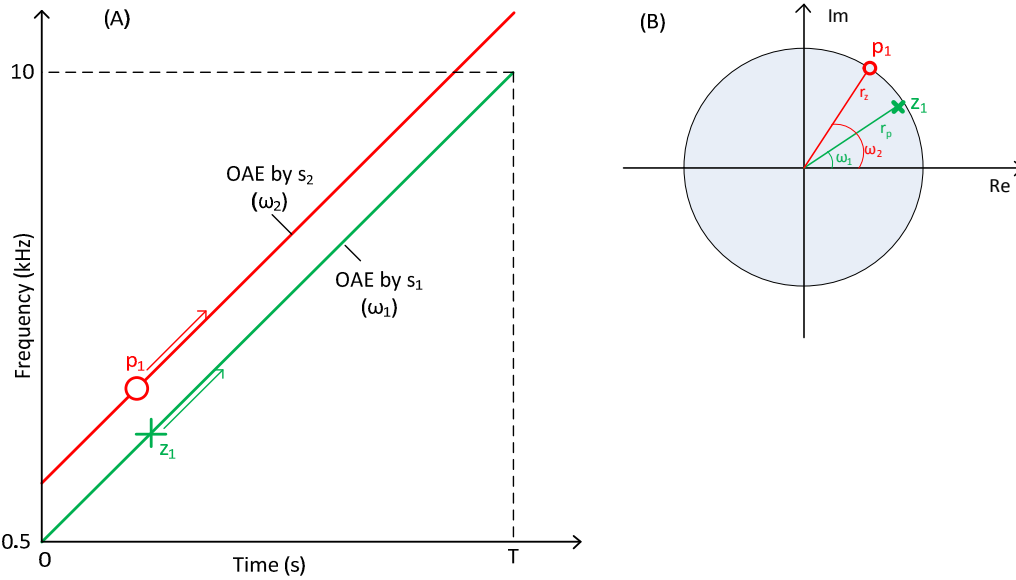$$G = \frac{1 - r_p}{1 - r_z \exp[j(\omega_2 - \omega_1)]} \tag{20}$$

Figure 31. The zeros and poles of the tracking filter.

The tracking filter was then applied to the temporal waveform of the residual

$[x(n), n = 0,1,2,\text{L } N-1]$ by the following difference equation:

$$y(n) = p_1 y(n-1) + G\left[x(n) - z_1 x(n-1)\right] \tag{21}$$

where y(n) is the filtered output which contains mainly the swept-tone SFOAE evoked by

s_1.

118

In this study, the bandwidth of the tracking filter Δf was chosen to the square root of the sweeping rate of the swept tone stimulus s_1:

$$\Delta f = \sqrt{(f_2 - f_1)/T} \tag{22}$$

Where f_1 is the starting frequency, f_2 is the ending frequency and T is the duration. Moreover, the filtered output $y(n)$ could be fed back to the same tracking filter several times in order to improve the selectivity of the tracking filter. Such operation was equivalent to increasing the order of the tracking filter and therefore could improve the performance.

In this study, the tracking filter technique was used to extract the swept-tone SFOAEs from the residuals. One way to demonstrate the effectiveness of the tracking filter was to analyze the spectrogram (a representation of energy distribution as a function of time and frequency) before and after the tracking filtering (Figure 32). As shown in panel (A), the spectrogram of the residual before the filtering showed two parallel lines, representing potential SFOAEs evoked by s_1 and s_2, respectively. Random noises were also distributed everywhere in the spectrogram. If such a residual was taken an FFT without any pre-processing, the amplitude spectrum [panel (C)] was rather noisy and no swept-tone SFOAEs could be clearly observed. In contrast, the spectrogram in panel (B) showed that only the s_1-evoked SFOAEs were left after the filtering and all other irrelevant noises were greatly attenuated. The corresponding amplitude spectrum [panel (D)] showed clear swept-tone SFOAEs.
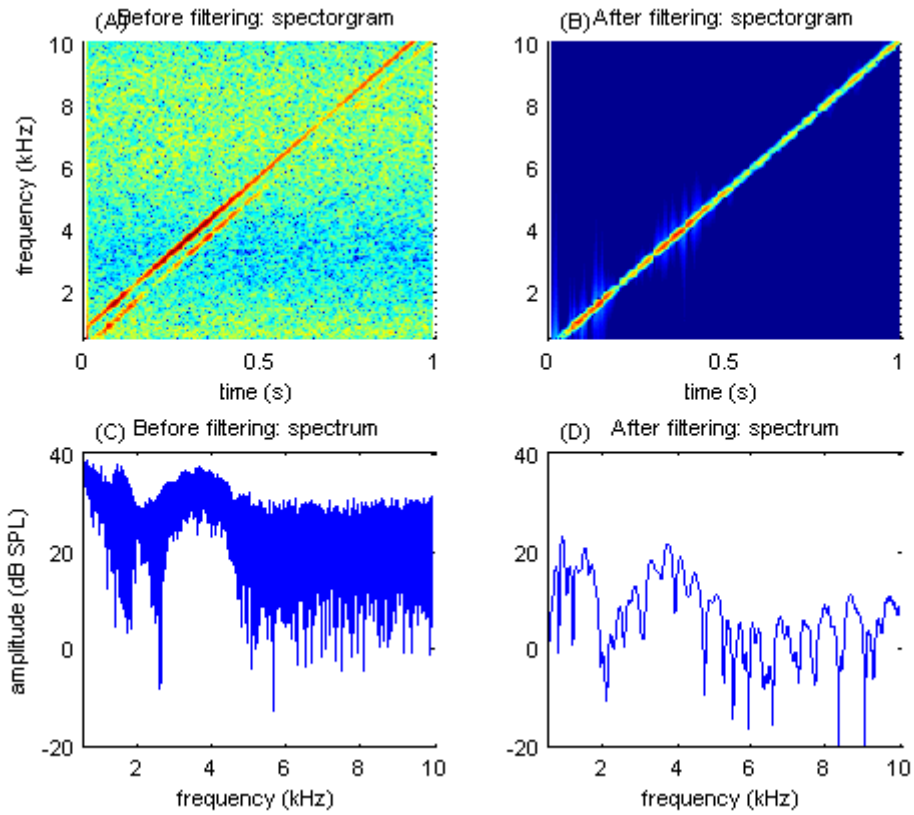
Figure 32. The spectrograms and amplitude spectra of the swept-tone SFOAEs before and after the tracking filtering.