Hearing Research 370 (2018) 201–208

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

No otoacoustic evidence for a peripheral basis of absolute pitch

Larissa McKetton [a], David Purcell [b], Victoria Stone [b], Jessica Grahn [c], Christopher Bergevin [d, *]

[a] Biology, York University, Toronto, ON, Canada
[b] Communication Sciences and Disorders, University of Western Ontario, London, ON, Canada
[c] Psychology, University of Western Ontario, London, ON, Canada
[d] Physics & Astronomy, York University, Toronto, ON, Canada

ARTICLE INFO

Article history:
Received 5 June 2018
Received in revised form
31 July 2018
Accepted 6 August 2018
Available online 8 August 2018

ABSTRACT

Absolute pitch (AP) is the ability to identify the perceived pitch of a sound without an external reference. Relatively rare, with an incidence of approximately 1/10,000, the mechanisms underlying AP are not well understood. This study examined otoacoustic emissions (OAEs) to determine if there is evidence of a peripheral (i.e., cochlear) basis for AP. Two OAE types were examined: spontaneous emissions (SOAEs) and stimulus-frequency emissions (SFOAEs). Our motivations to explore a peripheral foundation for AP were several-fold. First is the observation that pitch judgment accuracy has been reported to decrease with age due to age-dependent physiological changes cochlear biomechanics. Second is the notion that SOAEs, which are indirectly related to perception, could act as a fixed frequency reference. Third, SFOAE delays, which have been demonstrated to serve as a proxy measure for cochlear frequency selectivity, could indicate tuning differences between groups. These led us to the hypotheses that AP subjects would (relative to controls) exhibit a. greater SOAE activity and b. sharper cochlear tuning. To test these notions, measurements were made in normal-hearing control ($N = 33$) and AP-possessor ($N = 20$) populations. In short, no substantial difference in SOAE activity was found between groups, indicating no evidence for one or more strong SOAEs that could act as a fixed cue. SFOAE phase-gradient delays, measured at several different probe levels (20-50 dB SPL), also showed no significant differences between groups. This observation argues against sharper cochlear frequency selectivity in AP subjects. Taken together, these data support the prevailing view that AP mechanisms predominantly arise at a processing level in the central nervous system (CNS) at the brainstem or higher, not within the cochlea.

1. Introduction

1.1. Background- absolute pitch

Absolute pitch (AP) is the remarkable ability to discern the pitch of an acoustic stimulus in the absence of a reference (e.g., identify a note played on a piano when listening without any other visual or auditory cue). AP ability, useful for musicians, presumably affects other functional aspects of hearing such as speech recognition (Deutsch, 2013). The exact prevalence of AP in the general population is not well documented, however estimates of 1/10,000 have been reported (Profita et al., 1988). Opinions vary as to whether absolute pitch is genetic (Profita et al., 1988; Theusch et al., 2009; Theusch and Gitschier, 2011) or a learned ability that is linked to one's exposure to music up to a critical age (Deutsch et al., 2006; Zatorre, 2003). Supporting evidence for the latter suggests that an earlier onset of musical training in note-labeling up until the age of seven is linked with a higher probability of AP development (Deutsch et al., 2006; Miyazaki et al., 2012; Miyazaki and Ogawa, 2006).

Explanations for AP have primarily focused on neuroanatomical differences at the level of the central nervous system (CNS). For example, structural brain differences revealed an increased leftward asymmetry of the planum temporale volume in subjects with AP compared to controls (Keenan et al., 2001), and enhanced functionally connected networks at the cortical level (Loui et al., 2012; Schulze et al., 2009). Our motivation here considered

* Corresponding author. 4700 Keele St, Petrie 240, Toronto, ON, M3J. 1P3, Canada.
E-mail addresses: mcketton@yorku.ca (L. McKetton), purcelld@nca.uwo.ca (D. Purcell), vstone2@uwo.ca (V. Stone), jgrahn@uwo.ca (J. Grahn), cberge@yorku.ca (C. Bergevin).

https://doi.org/10.1016/j.heares.2018.08.001

whether cortical/cognitive differences were sufficient to explain AP development, or if there was a peripheral foundation (i.e., a cochlear basis) as evidenced by otoacoustic emissions (OAE) properties.

1.2. OAEs

OAEs are sounds produced by the inner ear and are measured using a sensitive microphone in the ear canal (Kemp, 1978; Shera and Abdala, 2012; Bergevin et al., 2017). During forward transduction, hair cells convert mechanical stimuli into electrical responses. In reverse transduction, the electrical responses of outer hair cells (OHCs) induce a mechanical output that enables them to act as force generators. As a result, the inner ear exhibits a nonlinear power amplification to boost the detection of low-level signals, as well as sharpen frequency selectivity. Healthy ears emit OAEs, typically thought to be a by-product of this amplification mechanism at work in the cochlea (Lonsbury-Martin and Martin, 2003).

Spontaneous otoacoustic emissions (SOAEs) are emitted in the absence of any external stimulus. They manifest as distinct low-level narrowband peaks in the spectral magnitude of the ear canal sound pressure. Their prevalence is roughly 40-60% in individuals who have normal audiological thresholds (Snihur and Hampson, 2011; Talmadge et al., 1993). The presence of SOAEs is usually considered to be a sign of normal cochlear health, but the absence of SOAEs is not necessarily an indication of abnormality. When present, SOAE peak frequencies are unique to each ear, akin to a fingerprint. Unless perturbed by moderate to high-level stimuli, human SOAEs tend to exhibit relatively low variability in frequency and magnitude. On time scales of ~1 min, half-power bandwidths were reported less than 0.1% of the SOAE frequency while rms magnitude fluctuations were reported less than 6.3% of the SOAE mean magnitude (van Dijk and Wit, 1990). SOAE frequencies and magnitudes have been reported to be more variable on the order of hours to months. Although SOAE peaks are relatively stable when subjects are seated quietly in a booth during a recording session, they can fluctuate in magnitude (average drifts on the order of 1.5 dB, typically increasing) and in frequency (between 0.1 and 0.5% in either direction) (Whitehead, 1991). Over long-time periods (e.g., years), SOAEs non-uniformly decrease in magnitude and uniformly decrease in frequency at 0.25%/year from just after birth to at least 60 years of age (Baiduc et al., 2014; Burns, 2009; Abdala et al., 2017).

SOAEs have been tied to perception in that they can interact with sound stimuli, which is sufficient to act as a detection cue (Long and Tubis, 1988; Zwicker and Schloth, 1984). For example, external tones presented at low-levels close to threshold interact with SOAE peaks in a qualitatively different fashion depending upon tone frequency: When nearby (but not atop of) an SOAE peak, a beating perception arises while when very close to the SOAE a more tonal sensation is perceived (Long and Tubis, 1988). Additionally, these studies found that subjects with SOAEs had lower auditory thresholds near the SOAE peak. It has also been found that the presence of SOAEs can affect measures of frequency selectivity via psychophysical tuning curves (PTCs) (Micheyl and Collet, 1994; Baiduc et al., 2014). Micheyl and Collet et al. reported that subjects with SOAEs had sharper PTCs at 2 kHz, but not at the other tested frequencies of 1 or 4 kHz. In addition, a previous study looked at the effect of SOAEs on PTCs by comparing tuning between an ear with SOAEs to the other ear without any SOAEs in an individual subject. The findings revealed that PTCs were sharper in the ear where SOAEs were present as compared to an equivalent frequency in the other non-emitting ear, or another non-SOAE frequency in the same ear (Bright, 1985; 2007).

Stimulus frequency otoacoustic emissions (SFOAEs) are evoked via an external stimulus and arise at the same frequency as the elicitor. One theory posits that they arise due to impedance irregularities within the cochlea that cause a coherent scattering of cochlear traveling waves (Zweig and Shera, 1995). SFOAEs provide a measure of mechanical delay within the cochlea (Goodman et al., 2004), and these delays have been shown to be directly related to the frequency selectivity of the ear (Joris et al., 2011; Kemp, 1986; Moleti and Sisto, 2016; Neely et al., 1988; Shera et al., 2002; Shera and Guinan, 2003; Bergevin and Shera, 2010). As such, SFOAEs have been used to systematically compare tuning estimates across different species (Bergevin et al., 2010, 2015; Bergevin et al., 2010).

1.3. Motivation

Previous studies have reported gradual decreases in pitch judgment accuracy in AP with age (Athos et al., 2007; Vernon, 1977). These authors hypothesized that age-related changes in the cochlear map (basal shift) may underlie the trend toward overestimation of pitch (sharpening) in those with AP over 50. These findings link errors in AP judgement to age-dependent physiological changes in mechanical events in the cochlea that may underlie this phenomenon.

The motivations for this study linking cochlear mechanics and AP included the concept that SOAEs are related to perception and could act as a spectral benchmark. This led us to hypothesize that AP subjects would exhibit greater SOAE activity relative to controls. Specifically, we predicted that we might see more SOAE peaks and/or the existence of relatively large SOAE peaks in AP subjects. We also hypothesized that relatively sharper peripheral tuning may provide a cue for AP. Support for this hypothesis stems from a study that found sharper cochlear tuning in a high-frequency 4 kHz region in musicians compared to non-musician controls, as measured by physiological tuning curves from SFOAEs and psychophysical tuning curves derived via simultaneous masking (Bidelman et al., 2016). Although their focus was not on AP, Bidelman et al. showed a relationship between the years of musical training and improved tuning in musicians. We therefore predicted that if AP possessors have sharper frequency tuning of their peripheral auditory filters, they would have longer SFOAE delays (Shera et al., 2002; Joris et al., 2011). Furthermore, we reasoned that if significant differences were not found, increased frequency selectivity relevant to AP could still occur, but not at the level of the cochlea. These results would be more in line with a study that found no significant differences between the just-noticeable difference (JND) thresholds of two-testing frequencies (1000 Hz and the equitempered tone of B5 987.76 Hz) in AP-musicians compared to non-AP musicians using a two-alternative forced choice task psychoacoustic experiment (McKetton, 2016). Additional support for no cochlear differences in AP stems from Bianchi et al. (2016) who found no peripheral frequency selectivity between musicians and non-musicians behaviorally and objectively for resolved and unresolved complex tones. An anecdotal point of motivation stemmed from the observation that in Asian populations, there is both increased incidence of AP (Deutsch et al., 1999; Gregersen et al., 2001) and SOAEs (Whitehead et al., 1993).

2. Methods

2.1. Data collection and participant information

We examined SOAEs and SFOAEs in both control ($N = 33$ subjects, 40 ears) and AP ($N = 20$, 30 ears) normal hearing populations, age range 18–48 years. Data were collected independently at Western University [WU; control ($N = 26$, 26 ears) and AP ($N = 9$, 9 ears)] and York University [YU; control ($N = 7$, 14 ears) and AP

($N = 11$, 21 ears)] using the same acquisition parameters and data analysis. Cumulatively from both testing locations, sixteen participants were excluded from OAE analyses due to either hearing loss, earwax build-up obstructing proper OAE measurement, or no usable data due to noisy measurements. Approval was given by the University Ethics Committee (YU) and the Health Sciences Research Ethics Board (WU). Prior to OAE testing, each participant gave written informed consent and filled out a questionnaire. Due to time constraints, the majority of WU subjects only had their dominant ear corresponding to their handedness tested. Of the WU subjects, 22/26 controls had their right ear tested, whereas 5/9 AP subjects had their right ear tested. York subjects usually had both ears tested. AP participants were recruited from notices at the university music departments and by word of mouth. More females were tested in each group (Table 1).

An AP test developed in the laboratory of Dr. Gottfried Schlaug was completed by all subjects to objectively classify AP status [http://www.musicianbrain.com/aptest/]. The AP test consisted of 24 sine wave tones taken from the chromatic scale (C4-B4 repeated twice and randomized per trial). Each subject had to name the note of the tone they heard by selecting one out of twelve possibilities. Data were collected on four trials (for a total of 96 tones presented). AP ability was confirmed if the accuracy was 90% or above on the responses given that were within one semitone of the presented tone pooled across the four trials (Hamilton et al., 2004; Miyazaki, 1988; Zatorre and Beckett, 1989). All subjects had normal audiometric thresholds that were under 20 dB HL for frequencies tested between 0.5 and 8 kHz using a 10 dB down, 5 dB up bracketing procedure measured for each participant using the Amplivox 240 (YU) or Madsen Itera (WU) diagnostic audiometer. During the OAE recording sessions, subjects remained awake and sat quietly in a double-walled acoustic sound-isolating chamber [Industrial Acoustics Co. (YU) or ECKOUSTIC (WU)].

OAEs were measured using an Etymotic ER-10C probe at YU, and an ER-10B+/ER2 system at WU. The ER-10C probe employed two stimulus transducers that generated the sounds, and a microphone that measured the stimuli and resulting OAEs. The resulting ear canal sound pressure was picked up by the microphone and was digitized using a soundcard (Lynx Two-A, Lynx studios). A 44.1 kHz sample rate at 24 bits/sample was used to process the data using custom software (Bergevin et al., 2015).

2.2. Experimental protocol - SOAE

Each subject waited fifteen minutes in the sound attenuated chamber before data acquisition to allow SOAE activity to stabilize. SOAE measurements were taken at the beginning and end of each experiment. An SOAE spectrum (Fig. 2B, black curve) was obtained by spectrally averaging the magnitudes of the Fourier transform of 60 artifact-free buffers of 32,768 points each (frequency bin width of 1.35 Hz). Classifying segments as “artifact-free” was achieved via subtraction of two successive waveforms, where the maximum value from the difference was taken and assessed if it was above a subject-specific threshold (see Shera and Guinan, 1999). If SOAE peaks were present, an additional 120 s waveform was taken to allow for more detailed analysis if required.

Quantitative systematic comparisons of SOAE properties were made between the two groups. These included pooled group population comparisons (Fig. 1) for the number of SOAE peaks, their magnitude, width, and “noisiness” (i.e., the signal-to-noise ratio, SNR). A custom-coded peak-picking algorithm was used to objectively identify SOAE peaks from an individual SOAE spectrum. These values were verified by visual inspection. To reduce spurious counts, only peaks whose magnitude was at least 3 dB above noise floor were included for further analysis. The SOAE magnitudes, width and SNR were computed via a Lorentzian fit at each peak (via nonlinear regression) in a localized neighborhood (± 20 Hz on either side). Peak width was determined as the full-width half-max (FWHM), whereas the SOAE peak SNR was determined as the dB difference between the peak maximum and the asymptotic limits.

For several subjects, their recorded spectra were partially contaminated by increased electrical line noise (60 Hz harmonic peaks). These noisy peaks were narrow, as their spectral energy was confined to a single frequency bin as compared to SOAE peaks, which spread across several frequency bins. Any peak that was a harmonic of 60 Hz was excluded unless they exhibited a spectral width consistent with SOAE peaks. Pearson chi-square analyses were conducted to determine if any differences in SOAE number, peak magnitude, peak width, and peak SNR were found between groups from pooled data from both testing sites [AP ($N = 20$, 30 ears), control ($N = 33$, 40 ears)], from one ear selected at random in the group that had both ears measured [AP ($N = 20$, 20 ears), control ($N = 33$, 33 ears)], and for equal group sample size comparisons [AP ($N = 20$, 20 ears), control ($N = 20$, 20 ears)]. Additionally, a one-way analysis of covariance (ANCOVA) was conducted to determine if there were any statistically significant differences between the number of SOAEs for each group controlling for sex and age.

2.3. Experimental protocol - SFOAE

Earphone calibration was performed in-situ using flat-spectrum noise and periodically checked throughout the experiment to ensure proper probe and suppressor levels. No corrections were made for ear canal standing waves in that we assumed the sound pressure measured at the probe tip was an indicator of sound pressure at the eardrum for frequencies below 10 kHz. SFOAEs were measured using a swept-tone paradigm (Kalluri and Shera, 2013) with the SFOAE extracted via suppression (Brass and Kemp, 1993; Kalluri and Shera, 2013; Neely et al., 2005; Shera and Guinan, 1999).

Table 1
Participant background information.

Group	AP			Controls		
	York U	WU	Total	York U	WU	Total
Number of subjects	11	9	**20**	7	26	**33**
Number of ears tested	21	9	**30**	13	26	**39**
Gender (male/female)	5/6	3/6	**8/12**	1/6	2/24	**3/30**
Age (years; Mean ± SD)	25 (9.5)	22.6 (4)	**24.5 (3.4)**	26 (4.5)	24.2 (3)	**24.1 (7.4)**
Handedness						
Right-handed	9	6	**15**	6	25	**31**
Left-handed	1	3	**4**	1	1	**2**
Ambidextrous	1	–	**1**	–	–	–
AP Test (percent; Mean ± SD)	98.5 (2.8)	100	**99.2 (2.2)**	14.2 (7.2)	18.2 (13.1)	**17.3 (12.8)**

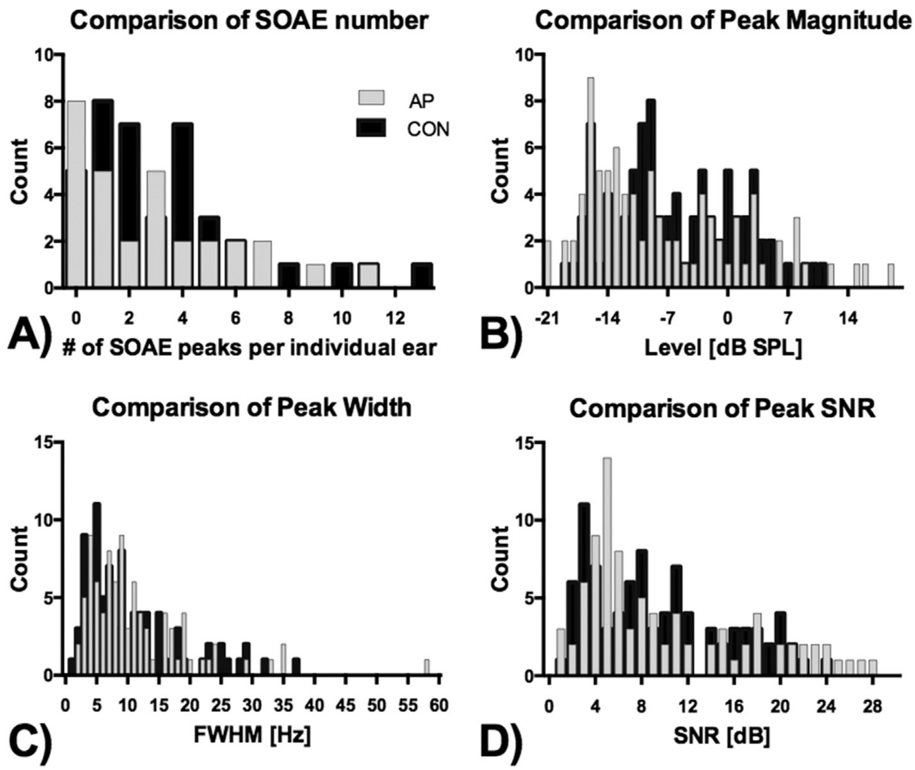

Fig. 1. SOAE peak histogram plots in pooled AP ($N = 20$, 30 ears) and control ($N = 33$, 39 ears) subjects for A) the number of SOAE peaks, B) SOAE magnitude (peak height) C) SOAE width (FWHM of a Lorentzian fit), and D) peak SOAE SNR. For each histogram comparison, Pearson chi-square analyses showed no significant differences between the two groups (all $p > 0.05$).

For the probe tone, the frequency range was 0.5–6 kHz and levels were examined from 20 to 50 dB SPL in 10 dB steps. The suppressor was fixed at 40 Hz above the probe tone and 15 dB higher in level. For averaging, 34 sweeps were presented, with a sweeping rate of 2 kHz/s.

Separation of SFOAE from the stimulus is a challenge because the emission, which is relatively small, occurs at the same frequency. To illustrate this visually, Fig. 2A shows the microphone response at the probe frequency when the tone was presented at a constant level ($L_p = 30$ dB SPL) for two conditions: probe alone (solid) and probe and suppressor (dashed). When the suppressor was also presented, it inhibited the generation of the SFOAE and caused the SFOAE measured in the ear canal to be reduced (or eliminated). The residual SFOAE (thick/solid trace in Fig. 2B) was then extracted by a least squares analysis window (Kalluri and Shera, 2013). The noise floor was determined from a single swept-tone measurement with a sweep rate of 1 Hz/ms and an analysis window of 25, 100 and 200 ms (Kalluri and Shera, 2013). The SOAE spectrum, obtained just prior to the SFOAE recording, is also shown (thin/solid trace in Fig. 2B).

We calculated SFOAE phase gradient delays, defined as the negative of the slope of the emission phase (in cycles) versus frequency from unwrapped phase responses (Fig. 2C). As described in a previous theoretical study (Shera and Bergevin, 2012), a peak-picking algorithm was employed for extraction of the delays to help avoid errors associated with computation of the slopes about phase discontinuities. In brief, data points at a given frequency were included only if they occurred close to a peak in emission magnitude, as well as exhibited at least a 10 dB SNR. As described in the Results section, other strategies were explored and yielded qualitatively similar results. Delays were subsequently expressed in dimensionless form as the equivalent number of stimulus periods (N_{SFOAE}; Shera et al., 2002) as shown in Fig. 2C. Fig. 2D shows the N_{SFOAE} extracted from the representative subject. For frequency regions with sufficient SNR, individual N_{SFOAE} curves are robust, in that the scatter apparent in the data is typical of SFOAEs and reproducible, and does not arise from pathology or measurement noise (Bergevin et al., 2012; Kalluri and Shera, 2013) (Fig. 2D). In fact, SFOAE responses and the associated phase trends are highly reproducible within a session (e.g., Bergevin et al., 2012). Rather than noise, individual trends presumably reflect the role of mechanical irregularity inherent in the process of emission generation (Shera and Guinan, 1999).

In order to compare N_{SFOAE} properties, and thereby tuning estimates across AP and control groups, trends were computed from grouped data using locally-weighted linear regression ("*loess*", Cleveland, 1993) (Fig. 3). Confidence intervals were then determined using bootstrap resampling taken 100 times (shaded regions in Fig. 3; Bergevin et al., 2010) to determine 95% confidence intervals. For this, the subject identifier was the key resampling parameter. That is, for each subject in each group (AP or control), an array of values was created. Those arrays were then resampled with replacement to create a bootstrapped ID array that was then used to create a pooled array of associated N_{SFOAE} values across frequency for which a trend was computed for each group. This process was then repeated such that uncertainty estimates of 95% confidence intervals in the trends could be determined. Additional SOAE and SFOAE analyses on subjects who had both ears measured were explored by selecting only one ear at random for inclusion to omit any biases for both SOAE and SFOAE analyses.

3. Results

Our results indicate no statistically significant differences in the

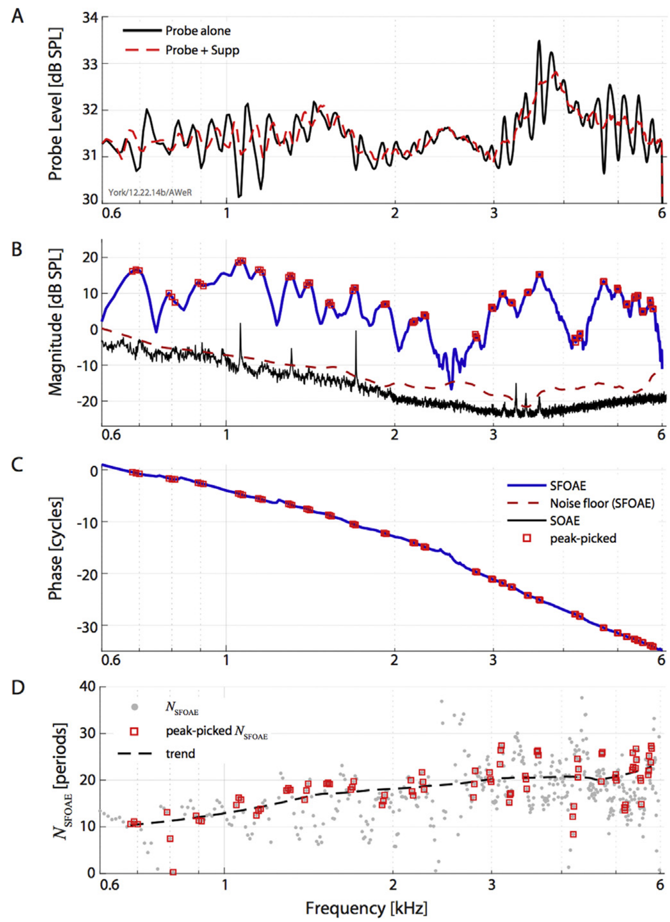

Fig. 2. Representative SFOAE and SOAE data from a single ear (Lp = 30 dB SPL). A) The probe-alone condition (solid line) depicts the measured pressure at the stimulus frequency as a combination of both the stimulus and the emission, whereas when the suppressor is also presented (dashed line), the interference at the probe frequency diminishes, indicating greater dominance of the stimulus. B) SFOAEs (top solid line curve) were extracted by a least squares analysis. The estimated noise floor for the SFOAE measurement (dashed line) is also included. Also shown (open gray squares) are the corresponding frequencies flagged via the "peak-picking" algorithm (Shera and Bergevin, 2012) used for extracting phase gradient delays. The SOAE curve (black line) is depicted under the SFOAE curve. C) The associated SFOAE phase (unwrapped). The slope of this phase curve with respect to frequency reveals the phase-gradient delay, which is the basis for N_{SFOAE} as shown in Fig. 3 and panel D of Fig. 2. D)N_{SFOAE} curve extracted from the representative subject denoting all points (filled gray circles), peak-picked points (open black squares), and trend line (dashed line).

number of SOAEs in AP compared to control subjects based on our pooled SOAE counts (AP: $N = 20$, 30 ears, control: $N = 33$, 38 ears) subjects χ^2 (3) = 1.8, $p = 0.62$ (Fig. 1A). In addition, there were no significant differences for peak magnitudes χ^2 (5) = 6.04, $p = 0.3$ (Fig. 1B), peak widths χ^2 (2) = 0.05, $p = 0.97$ (Fig. 1C), and peak SNR χ^2 (5) = 8.48, $p = 0.13$ (Fig. 1D) between groups.

Additional analyses were conducted to determine if there were any group differences in the SOAE peak incidence, magnitude, width, and SNR for the number of randomly selected ears (bootstrapped) for equal group sample size comparisons and for comparisons between testing sites (i.e., randomly selecting 1 ear of those who had 2 ears tested, and by comparing the same sample size between groups). Pearson chi-square analyses revealed no significant differences between groups when accounting for equal group comparison and testing location for each condition (all $p > 0.05$). Lack of difference in SOAE characteristics suggests that AP does not depend on the use of an SOAE as an internal reference tone. A one-way ANCOVA showed no significant differences in the

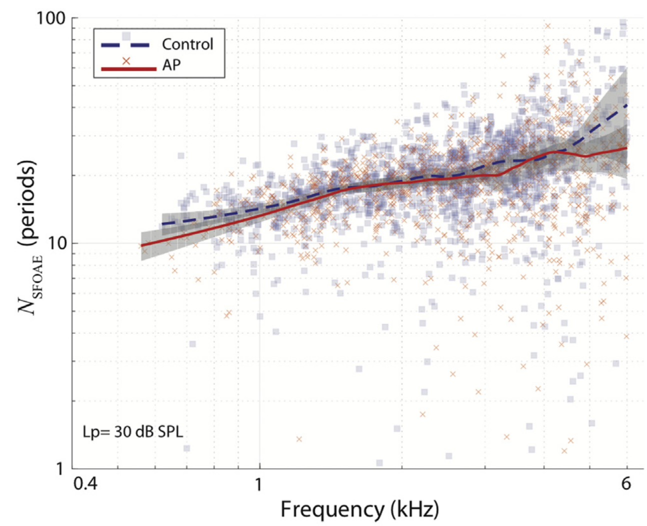

Fig. 3. Comparison of pooled SFOAE phase-gradient delays (as number of stimulus periods, N_{SFOAE}) for both AP (x symbols and solid line; $N = 20$ individual ears) and control (square symbols and dashed line, $N = 33$) groups. Probe level used here was $L_p = 30$ dB SPL. N_{SFOAE} values were extracted using a peak-picking routine (Shera and Bergevin, 2012). Shaded areas indicate a 95% confidence interval (CI) via bootstrapping across subjects. Results suggest no overall improved frequency selectivity in AP subjects. Similar relationship exists at other stimulus levels, although delays were progressively longer for lower stimulus levels (see Results).

number of SFOAEs in each group F (52) = 0.037, $p = 0.85$; no significant sex differences between each group F (52) = 0.043, $p = 0.84$; and no significant differences in age between groups F (52) = 1.31, $p = 0.26$.

Although measurements were made at a variety of probe levels, for SFOAEs we focus here on $L_p = 30$ dB SPL, which yields a reasonable compromise between a "low level" stimulus (Bergevin et al., 2015) and a sufficient SNR. SFOAE delays (in dimensionless form, N_{SFOAE}), serving here as a proxy measure of peripheral tuning, are shown in Fig. 3. For this plot, subjects were pooled together from both locations for each group (e.g., for controls, 7 ears from York and 26 from WU, for a total of $N = 33$). To make comparisons of tuning estimate based upon N_{SFOAE} between AP and control groups, trend lines were computed. Shaded areas indicate a 95% confidence interval (CI) via bootstrapping across subjects (see Methods). Overall, the pooled N_{SFOAE} results were similar between both AP and control groups in that they overlapped in trends and/or uncertainty estimates, irrespective of frequency.

We systematically explored a range of analysis approaches for N_{SFOAE} comparison between groups, beyond that made in Fig. 3. The purpose was to rule out any sort of systematic error when comparing the relevant trends and uncertainties. First, similar overlap in trends was observed when comparing AP and control subjects from just the WU measurement location or just the York location. Thus a similar result was obtained for two different populations measured at two different locations. Second, we examined the effect of stimulus level (L_p). To a rough first order based upon the trends, N_{SFOAE} increased by 1.15–1.2 for every 10 dB decrease in L_p at 1 kHz and 1.25–1.3 at 3 kHz. This was true for both (pooled) AP and control groups. For a given L_p (20, 30, 40, and 50 dB SPL were examined), when comparing groups, the trends (and uncertainties) generally overlapped (similar to that shown in Fig. 3). There was some variability in the trends across frequency (especially for $L_p = 20$ dB SPL), but no systematic differences were readily apparent between the two groups. Third, (for $L_p = 30$ dB SPL), we compared determination of N_{SFOAE} and the associated trends using a peak-picking algorithm (Shera and Bergevin, 2012) versus a simple inclusion criterion based upon a 10 dB SNR, as employed in previous studies (e.g., Bergevin et al., 2010; Bergevin et al., 2010). We observed that peak-picking generally caused N_{SFOAE} trends to increase slightly (especially between 1.5 and 3.5 kHz), and that this effect was similar for both groups. Thus, the similarity in AP and control N_{SFOAE} trends is independent of the delay extraction method. Fourth, we examined whether having multiple SOAE peaks introduced any sort of bias. Consistent with the results of Bergevin et al. (2012), N_{SFOAE} trends were similar between those with and without SOAE activity. When comparing AP and control groups with the inclusion criteria that a subject had to have at least two or more SOAE peaks, the N_{SFOAE} trends were indistinguishable up to 5 kHz. Above that, the control group exhibited larger N_{SFOAE} values for the sample population measured. Lastly, we examined bootstrapping the pooled data rather than using individual subjects, to see how estimates of uncertainty were affected. The main effect was to reduce the uncertainty slightly (i.e., narrower error bars in Fig. 3). However, the overlap remained for all frequencies below 5 kHz. Taken together, we argue that the similarity shown in Fig. 3 is robust and not subject to systematic error stemming from the analysis method.

Additionally, we compared SFOAE delays to previous studies. Direct comparisons are complicated since SFOAE delays depend strongly upon level. Schairer et al. (2006) examined a variety of probe levels, but only at 40 dB SPL and higher (Schairer et al., 2006). Shera et al. (2002) reported values only for a probe level of 40 dB SPL. Bergevin et al. (2012) reported delays for 20 dB SPL. Dewey and Dhar (2017) used a 36 dB level, but they calibrated differently in that they used forward pressure level (FPL) rather than dB SPL at the probe tip, making it harder to compare across studies (Dewey and Dhar, 2017). Upon doing so, they get values consistent with Shera et al., 2002, that were 40 dB SPL at probe tip. However, none of these studies used 30 dB SPL, as focused on in Fig. 3. To facilitate comparisons, Fig. 4 pools together all SFOAE delays measured across different levels examined in this study, and includes the 40 dB SPL probe level delays from Shera et al. (2002). A linear ordinate is used here to improve visual comparison. Overall, we saw a systematic increase in delay with decreasing probe level,

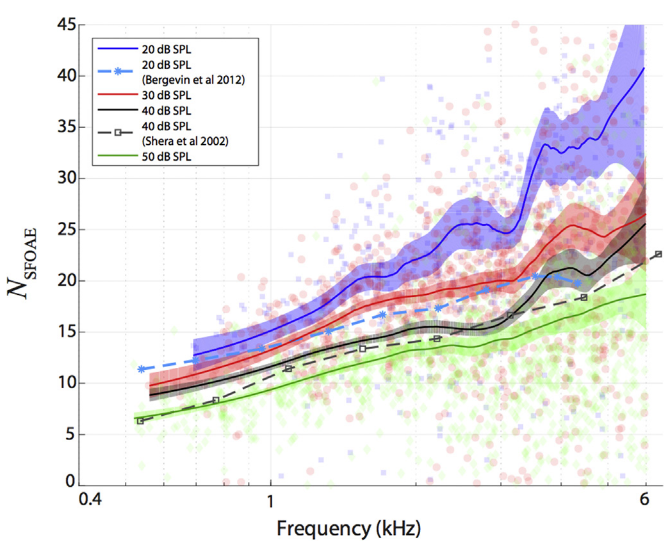

Fig. 4. Level-dependence of SFOAE delays. All subjects were pooled together and one ear per subject was included. A peak-picking paradigm was employed. The probe level is indicated in the legend, and the suppressor level used was 15 dB higher in all cases (i.e., Ls-Lp was constant at 15 dB). Also included are trend lines from two previously published reports.

consistent with the above-mentioned studies.

The data from Fig. 4 on SFOAE delays compared to different studies showed that at 1 kHz, the delay decreased by 24–29% for a 20 dB increase in probe level, whereas at 2 kHz, changes were closer to 28-31%. These values are smaller than the 38% change for ABR latencies noted by Neely and Rasetshwane (2017) (see also Neely et al., 1988). This discrepancy may be related to differences in generation mechanisms where SFOAEs may not be affected by effects such as "synaptic adaptation". Further, the delays reported here were generally the same or larger than those previously reported. For example, the SFOAE delays at 30 dB SPL in our study were larger than those reported by Bergevin et al. (2012) at 20 dB SPL (Fig. 4). The reason for this discrepancy is unclear, though the previous studies included in Fig. 4 used a discrete tone paradigm, whereas our study used a swept tone stimulus with a different sample size. Nevertheless, we do not expect differences in delay across studies to confound comparison between control and AP groups within the current study.

4. Discussion

This study investigated whether there was evidence for differences in cochlear function between individuals with and without AP. Our results revealed no otoacoustic-based difference in cochlear function for AP using SOAEs and SFOAEs. Based on our pooled SFOAE phase-gradient delays, both AP and control data showed sufficient overlap below 5 kHz that the trends were indistinguishable (Fig. 3). Around 4 kHz and above, the groups diverged with the control group having numerically larger N_{SFOAE} values compared with the AP group, although the groups were not statistically significantly different. Our findings deviate from Bidelman et al. (2016) that found sharper tuning at 4 kHz in musicians using physiological tuning curves from SFOAEs and with psychophysical tuning curves derived via simultaneous masking. It may be that the different methods yielded different results. Bidelman et al. used SFOAE suppression tuning curves that reveal only an indirect measure of cochlear tuning, whereas we estimated SFOAE tuning based on group delay that has previously been shown to have good agreement with auditory nerve responses (Shera et al., 2010). In addition, our results are in line with findings that found no significant differences between the JND thresholds in AP-musicians compared to non-AP musicians (McKetton, 2016). These findings suggest that AP ability arises primarily via central processing rather than specialized peripheral encoding.

A number of scientific studies have examined neuroanatomical differences in AP. For example, at the level of the CNS, structural findings revealed an increased leftward asymmetry of the planum temporale volume in subjects with AP compared to controls (Keenan et al., 2001), and enhanced functional networks at the cortical level (Loui et al., 2012; Schulze et al., 2009). Furthermore, AP possessors were found to have an enhanced auditory digit span (i.e., auditory working memory for numbers) compared to matched non-AP musicians, signifying that auditory working memory may play a fundamental role in AP emergence (Deutsch and Dooley, 2013). A number of studies have looked at AP in infants with varying results. It was previously reported that in a tone-sequence statistical learning task, infants were more likely to track patterns of absolute pitches rather than relative pitch representations. In contrast, adults relied more on relative pitch cues based on tone sequences (Saffran and Griepentrog, 2001). This finding suggested a shift during development from absolute to relative pitch processing that may be useful in speech perception. However, a later study found that infants preferred a novel, rather than a familiar, melody regardless of transposition and showed no preference for a familiar melody transposed to a novel pitch compared to the original pitch version (Plantinga and Trainor, 2005). This suggests that infants respond to relative but not absolute pitch cues.

4.1. Limitations

It is possible that our method to detect SFOAE cochlear tuning differences may not be sensitive enough. For example, differences between AP and control groups may exist but are relatively small and drowned out either in measurement noise or the analysis methodology comparing gross trends in N_{SFOAE}. Furthermore, if such differences do exist but are small, they may be amplified as information ascends the CNS. Additionally, our approach focused on the frequency range of 1–5 kHz, where SOAE activity predominantly lies and SFOAEs are relatively easy to measure since noise floors increase substantially for lower frequencies. Our current approach would not detect differences in peripheral function that are specific to frequencies <1 kHz.

Another possible limitation in our study was not controlling for musicianship. This would have been achieved by including a separate group that had musical experience without AP. While we did not account for musicianship, other studies found no significant differences between musicians and non-musicians in the mean evoked OAE magnitude (Perrot et al., 1999), or in peripheral differences in pitch discrimination thresholds for harmonic complex tones between 100 and 500 Hz (Bianchi et al., 2016). Overall, we argue that the methodological approach taken here has demonstrated that there is no obvious evidence to argue for differences in cochlear processing between AP and control groups.

Acknowledgments

Early support from Prof. Dorothy deVal (York) is acknowledged. Financial support came from the Natural Sciences and Engineering Research Council of Canada (CB, DP), (RGPIN - 430761-2013, RGPIN-05320-2007) and the Canada Foundation for Innovation (DP) (11419).

References

Abdala, C., Luo, P., Shera, C.A., 2017. Characterizing spontaneous otoacoustic emissions across the human lifespan. J. Acoust. Soc. Am. 141, 1874–1886. https://doi:10.1121/1.4977192.

Athos, E.A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., Gitschier, J., 2007. Dichotomy and perceptual distortions in absolute pitch ability. Proc. Natl. Acad. Sci. U.S.A. 104, 14795–14800. https://doi.org/10.1073/pnas.0703868104.

Baiduc, R.R., Lee, J., Dhar, S., 2014. Spontaneous otoacoustic emissions, threshold microstructure, and psychophysical tuning over a wide frequency range in humans. J. Acoust. Soc. Am. 135, 300–314. https://doi.org/10.1121/1.4840775.

Bergevin, C., Shera, C.A., 2010. Coherent reflection without traveling waves: On the origin of long-latency otoacoustic emissions in lizards. J. Acoust. Soc. Am. 127 (4), 2398–2409.

Bergevin, C., Fulcher, A., Richmond, S., Velenovsky, D., Lee, J., 2012. Interrelationships between spontaneous and low-level stimulus-frequency otoacoustic emissions in humans. Hear. Res. 285, 20–28.

Bergevin, C., Manley, G.A., Köppl, C., 2015. Salient features of otoacoustic emissions are common across tetrapod groups and suggest shared properties of generation mechanisms. Proc. Natl. Acad. Sci. U.S.A. 112, 3362–3367. https://doi.org/10.1073/pnas.1418569112.

Bergevin, C., Velenovsky, D.S., Bonine, K.E., 2010. Tectorial membrane morphological variation: effects upon stimulus frequency otoacoustic emissions. Biophys. J. 99, 1064–1072. https://doi.org/10.1016/j.bpj.2010.06.012.

Bergevin, C., Verhulst, S., van Dijk, P., 2017. Remote Sensing the Cochlea: Otoacoustics. In: Understanding the Cochlea. Springer, Cham, pp. 287–318.

Bianchi, F., Santurette, S., Wendt, D., Dau, T., 2016. Pitch discrimination in musicians and non-musicians: effects of harmonic resolvability and processing effort. J. Assoc. Res. Otolaryngol. 17, 69–79. https://doi.org/10.1007/s10162-015-0548-2.

Bidelman, G.M., Nelms, C., Bhagat, S.P., 2016. Musical experience sharpens human cochlear tuning. Hear. Res. 335, 40–46. https://doi.org/10.1016/j.heares.2016.02.012.

Brass, D., Kemp, D.T., 1993. Suppression of stimulus frequency otoacoustic emissions. J. Acoust. Soc. Am. 93, 920–939.

Bright, K.E., 1985. Microstructure Audiograms and Psychophysical Tuning Curves from Subjects with Spontaneous Otoacoustic Emissions (Doctoral dissertation). Retrieved from https://arizona.openrepository.com/bitstream/handle/10150/187995/azu_td_8522805_sip1_m.pdf?sequence=1.

Bright, K.E., 2007. Spontaneous Otoacoustic Emissions in Populations with Normal Hearing Sensitivity. Otoacoustic Emissions. Clinical Applications, pp. 69–86.

Burns, E.M., 2009. Long-term stability of spontaneous otoacoustic emissions. J. Acoust. Soc. Am. 125, 3166–3176. https://doi.org/10.1121/1.3097768.

Cleveland, W.S., 1993. Visualizing Data. Hobart Press.

Deutsch, D., 2013. Absolute pitch. In: The Psychology of Music. Elsevier Inc., San Diego, pp. 141–182. https://doi.org/10.1016/B978-0-12-381460-9.00005-5.

Deutsch, D., Dooley, K., 2013. Absolute pitch is associated with a large auditory digit span: a clue to its genesis. J. Acoust. Soc. Am. 133, 1859–1861. https://doi.org/10.1121/1.4792217.

Deutsch, D., Henthorn, T., Dolson, M., 1999. Absolute pitch is demonstrated in speakers of tone languages. J. Acoust. Soc. Am. 106, 2267. https://doi.org/10.1121/1.427738.

Deutsch, D., Henthorn, T., Marvin, E., Xu, H., 2006. Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. J. Acoust. Soc. Am. 119, 719–722. https://doi.org/10.1121/1.2151799.

Dewey, J.B., Dhar, S., 2017. Profiles of stimulus-frequency otoacoustic emissions from 0.5 to 20 kHz in humans. J. Assoc. Res. Otolaryngol. 18, 89–110. https://doi.org/10.1007/s10162-016-0588-2.

Goodman, S.S., Withnell, R.H., De Boer, E., Lilly, D.J., Nuttall, A.L., 2004. Cochlear delays measured with amplitude-modulated tone-burst-evoked OAEs. Hear. Res. 188, 57–69. https://doi.org/10.1016/S0378-5955(03)00375-7.

Gregersen, P.K., Kowalsky, E., Kohn, N., Marvin, E.W., 2001. Early childhood music education and predisposition to absolute pitch: teasing apart genes and environment. Am. J. Med. Genet. 98, 280–282.

Hamilton, R.H., Pascual-Leone, A., Schlaug, G., 2004. Absolute pitch in blind musicians. Neuroreport 15, 803–806. https://doi.org/10.1097/00001756-200404090-00012.

Joris, P.X., Bergevin, C., Kalluri, R., Mc Laughlin, M., Michelet, P., van der Heijden, M., Shera, C.A., 2011. Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proc. Natl. Acad. Sci. U.S.A. 108, 17516–17520. https://doi.org/10.1073/pnas.1105867108.

Kalluri, R., Shera, C.A., 2013. Measuring stimulus-frequency otoacoustic emissions using swept tones. J. Acoust. Soc. Am. 134, 356–368. https://doi.org/10.1121/1.4807505.

Keenan, J.P., Thangaraj, V., Halpern, A.R., Schlaug, G., 2001. Absolute pitch and planum temporale. Neuroimage 14, 1402–1408. https://doi.org/10.1006/nimg.2001.0925.

Kemp, D.T., 1986. Otoacoustic emissions, travelling waves and cochlear mechanisms. Hear. Res. 22, 95–104.

Kemp, D.T., 1978. Stimulated acoustic emissions from within the human auditory system. J. Acoust. Soc. Am. 64, 1386–1391.

Long, G.R., Tubis, A., 1988. Investigations into the nature of the association between threshold microstructure and otoacoustic emissions. Hear. Res. 36, 125–138. https://doi.org/10.1016/0378-5955(88)90055-X.

Lonsbury-Martin, B.L., Martin, G.K., 2003. Otoacoustic emissions. Curr. Opin. Otolaryngol. Head Neck Surg. 11, 361–366.

Loui, P., Zamm, A., Schlaug, G., 2012. Enhanced functional networks in absolute pitch. Neuroimage 63, 632–640. https://doi.org/10.1016/j.neuroimage.2012.07.030.

Micheyl, C., Collet, L., 1994. Interrelations between psychoacoustical tuning curves and spontaneous and evoked otoacoustic emissions. Scand. Audiol. 23, 171–178.

Miyazaki, K., 1988. Musical pitch identification by absolute pitch possessors. Percept. Psychophysiology 44, 501–512.

Miyazaki, K., Makomaska, S., Rakowski, A., 2012. Prevalence of absolute pitch: a comparison between Japanese and Polish music students. J. Acoust. Soc. Am. 132, 3484–3493. https://doi.org/10.1121/1.4756956.

Miyazaki, K., Ogawa, Y., 2006. Learning absolute pitch by children. Music Percept. 24, 63–78. https://doi.org/10.1525/mp.2006.24.1.63.

Mcketton, L.A., 2016. Peripheral and Central Auditory Processing in People With Absolute Pitch (Doctoral dissertation). Retrieved from http://yorkspace.library.yorku.ca/xmlui/bitstream/handle/10315/32750/McKetton_Larissa_2016_PhD.pdf?sequence=2&isAllowed=y.

Moleti, A., Sisto, R., 2016. Estimating cochlear tuning dependence on stimulus level and frequency from the delay of otoacoustic emissions. J. Acoust. Soc. Am. 140, 945. https://doi.org/10.1121/1.4960588.

Neely, S.T., Johnson, T.A., Garner, C.A., Gorga, M.P., 2005. Stimulus-frequency otoacoustic emissions measured with amplitude-modulated suppressor tones (L). J. Acoust. Soc. Am. 118, 2124–2127.

Neely, S.T., Norton, S.J., Gorga, M.P., Jesteadt, W., 1988. Latency of auditory brainstem responses and otoacoustic emissions using tone-burst stimuli. J. Acoust. Soc. Am. 83, 652–656.

Neely, S.T., Rasetshwane, D.M., 2017. Modeling signal propagation in the human cochlea. J. Acoust. Soc. Am. 142, 2155–2167. https://doi.org/10.1121/1.5007719.

Perrot, X., Micheyl, C., Khalfa, S., Collet, L., 1999. Stronger bilateral efferent influences on cochlear biomechanical activity in musicians than in non-musicians. Neurosci. Lett. 262, 167–170.

Plantinga, J., Trainor, L.J., 2005. Memory for melody: infants use a relative pitch code. Cognition 98, 1–11. https://doi.org/10.1016/j.cognition.2004.09.008.

Profita, J., Bidder, T.G., Optiz, J.M., Reynolds, J.F., 1988. Perfect pitch. Am. J. Med.

Genet. 29, 763–771. https://doi.org/10.1002/ajmg.1320290405.

Saffran, J.R., Griepentrog, G.J., 2001. Absolute pitch in infant auditory learning: evidence for developmental reorganization. Dev. Psychol. 37, 74–85.

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. J. Acoust. Soc. Am. 120, 901–914.

Schulze, K., Gaab, N., Schlaug, G., 2009. Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. BMC Neurosci. 10, 106. https://doi.org/10.1186/1471-2202-10-106.

Shera, C.A., Abdala, C., 2012. Otoacoustic Emissions: Mechanisms and Applications. Translational Perspectives in Auditory Neurocience: Hearing Across the Life Span-Assessment and Disorders, pp. 123–159.

Shera, C.A., Bergevin, C., 2012. Obtaining reliable phase-gradient delays from otoacoustic emission data. J. Acoust. Soc. Am. 132, 927–943. https://doi.org/10.1121/1.4730916.

Shera, C.A., Guinan, J.J., 2003. Stimulus-frequency-emission group delay: a test of coherent reflection filtering and a window on cochlear tuning. J. Acoust. Soc. Am. 113, 2762–2772.

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, 782–798.

Shera, C.A., Guinan, J.J., Oxenham, A.J., 2010. Otoacoustic estimation of cochlear tuning: validation in the Chinchilla. J. Assoc. Res. Otolaryngol. 11, 343–365. https://doi.org/10.1007/s10162-010-0217-4.

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, 3318–3323. https://doi.org/10.1073/pnas.032675099.

Snihur, A.W.K., Hampson, E., 2011. Sex and ear differences in spontaneous and click-evoked otoacoustic emissions in young adults. Brain Cognit. 77, 40–47. https://doi.org/10.1016/j.bandc.2011.06.004.

Talmadge, C.L., Long, G.R., Murphy, W.J., Tubis, A., 1993. New off-line method for detecting spontaneous otoacoustic emissions in human subjects. Hear. Res. 71, 170–182.

Theusch, E., Basu, A., Gitschier, J., 2009. Genome-wide study of families with absolute pitch reveals linkage to 8q24.21 and locus heterogeneity. Am. J. Hum. Genet. 85, 112–119. https://doi.org/10.1016/j.ajhg.2009.06.010.

Theusch, E., Gitschier, J., 2011. Absolute pitch twin study and segregation analysis. Twin Res. Hum. Genet. 14, 173–178. https://doi.org/10.1375/twin.14.2.173.

van Dijk, P., Wit, H.P., 1990. Amplitude and frequency fluctuations of spontaneous otoacoustic emissions. J. Acoust. Soc. Am. 88, 1779–1793.

Vernon, P.E., 1977. Absolute pitch: a case study. Br. J. Psychol. 68, 485–489.

Whitehead, M.L., 1991. Slow variations of the amplitude and frequency of spontaneous otoacoustic emissions. Hear. Res. 53, 269–280.

Whitehead, M.L., Kamal, N., Lonsbury-Martin, B.L., Martin, G.K., 1993. Spontaneous otoacoustic emissions in different racial groups. Scand. Audiol. 22, 3–10.

Zatorre, R.J., 2003. Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. Nat. Neurosci. 6, 692–695. https://doi.org/10.1038/nn1085.

Zatorre, R.J., Beckett, C., 1989. Multiple coding strategies in the retention of musical tones by possessors of absolute pitch. Mem. Cognit. 17, 582–589. https://doi.org/10.3758/BF03197081.

Zweig, G., Shera, C.A., 1995. The origin of periodicity in the spectrum of evoked otoacoustic emissions. J. Acoust. Soc. Am. 98, 2018–2047.

Zwicker, E., Schloth, E., 1984. Interrelation of different oto-acoustic emissions. J. Acoust. Soc. Am. 75, 1148–1154.