Time–frequency analysis of linear and nonlinear otoacoustic emissions and removal of a short-latency stimulus artifact

W. Wiktor Jedrzejczak
Institute of Physiology and Pathology of Hearing, ul. Zgrupowania AK “Kampinos” 1, 01-943 Warszawa, Poland

Andrew Bell
Research School of Biology, College of Medicine, Biology and Environment, The Australian National University, Canberra, ACT 0200, Australia

Piotr H. Skarzynski
Institute of Sensory Organs, ul. Mokra 1, Kajetany 05-830 Nadarzyn, Poland

Krzysztof Kochanek and Henryk Skarzynski
Institute of Physiology and Pathology of Hearing, ul. Zgrupowania AK “Kampinos” 1, 01-943 Warszawa, Poland

(Received 29 September 2011; revised 5 January 2012; accepted 9 January 2012)

Click-evoked otoacoustic emissions (CEOAEs) are commonly recorded as average responses to a repetitive click stimulus. If the click train has constant polarity, a linear average results; if it contains a sequence of clicks of differing polarity and amplitude, a nonlinear average can be calculated. The purpose of this study was to record both protocols from the same set of ears and characterize the differences between them. The major features of CEOAEs were similar under both protocols with the exception of a region spanning 0–5 ms in time and 0–2.2 kHz in frequency. It was assumed that the signal derived from the linear protocol was contaminated by stimulus artifact, and so a simple procedure was used—involving high-pass filtering and time-windowing—to remove components of this artifact. This procedure preserved the short-latency, high-frequency responses; it also produced a marked similarity in the time–frequency plots of recordings made under the two protocols. This result means it is possible to take advantage of the better signal-to-noise ratio of the linear data compared to its nonlinear counterpart. Additionally, it was shown that CEOAEs recorded under the linear protocol appear to be less dependent on the presence of spontaneous otoacoustic emissions (SOAEs).


PACS number(s): 43.64.Jb, 43.60.Hj [BLM] Pages: 2200–2208

I. INTRODUCTION

The most common method of recording click-evoked otoacoustic emissions (CEOAEs) uses a nonlinear protocol (Kemp et al., 1986). This method uses a train of three clicks of one polarity followed by a fourth having opposite polarity and three times greater amplitude. When the four otoacoustic responses are summed together, the result cancels out linear artifacts related to the acoustic stimulus (reflections from the walls of the external auditory canal and tympanic membrane) that tend to contaminate the early part of the cochlear echo. For this reason, the nonlinear protocol is the “gold standard” when using OAEs for diagnostic purposes. The linear protocol—where all stimuli are presented at the same level and polarity—has reduced diagnostic power (e.g., Moleti et al., 2002) and is used mainly in research settings (e.g., Yoshikawa et al., 2000; Hoth et al., 2007). It is generally accepted that the stimulus artifact, which affects the beginning of the CEOAE record under the linear protocol, hides the high-frequency, short-latency part of the response.

The stimulus artifact begins at the onset of the stimulus and continues until it dies out some 5 or 6 ms later (Ravazzani et al., 1996; Hoth et al., 2007). The problem is worst at high stimulus levels, and a number of studies has shown how CEOAEs recorded in the linear protocol are clearly contaminated by stimulus artifact at stimulus levels above about 65 dB SPL (e.g., Ravazzani et al., 1996; Hoth et al., 2007). The artifact can be removed by deleting all signal from this 5-ms-wide window, but then much of the high-frequency response, which has average latencies of 4–5 ms (Jedrzejczak et al., 2009a) would also be deleted. In comparison, if the nonlinear protocol is used, the excised stimulus window needs to be restricted to only 0–2.5 ms (Kemp et al., 1986).

Another important consideration is that for the same stimulus level, the linear protocol delivers greater cochlear responses, meaning that the signal-to-noise ratio and reproducibility are significantly better (Ravazzani et al., 1996; Tognola et al., 2001a; Hatzopoulos et al., 2003). Put another way, for the same signal-to-noise ratio, the linear protocol

a)Author to whom correspondence should be addressed. Electronic mail: w.jedrzejczak@ifps.org.pl

Copyright 2012 Acoustical Society of America
requires lower stimulus levels (Yoshikawa et al., 2000; Hatzopoulos et al., 2003) even though it calls for a wider initial window for artifact removal (Tognola et al., 2001a; von Specht et al., 2001). When studying CEOAEs from newborns, Tognola et al. (2001a) found that the linear protocol generated more “passes” (successful tests on a subject according to the ILO equipment’s internal reproducibility criterion) than the nonlinear protocol.

There have only been a few studies comparing CEOAEs evoked by both linear and nonlinear protocols (Ravazzani et al., 1996; Hoth et al., 2007). In general, the findings have been that both protocols give similar response profiles for stimulus levels lower than 65 dB SPL. With higher levels, artifacts increase until at 80 dB SPL, the stimulus artifact dominates the initial part of the waveform. The reported length of the stimulus artifacts generally ranges from 3.5 ms (Hatzopoulos et al., 2000a) to 6 ms (Tognola et al., 2001a), although one report suggested that only the response after 10 ms should be used (von Specht et al., 2001).

Given the linear protocol’s higher signal-to-noise ratio, the aim of the present study was to discover if a way could be found to identify the stimulus artifact more precisely and remove it from the linear OAE response. Of course, it is acknowledged that there are more artifacts present in CEOAEs than those directly due to the stimulus. For example, there are artifacts due to middle ear muscle reflexes and crossed olivocochlear bundle reflexes (Guinan et al., 2003; Zhao and Dhar, 2011). However, these artifacts are also present in CEOAEs recorded in the nonlinear protocol.

Time–frequency (TF) methods have been widely used in evaluating CEOAEs recorded under the nonlinear protocol (e.g., Wit et al., 1994; Jedrzejczak et al., 2009a; Hatzopoulos et al., 2009). These methods were used in theoretical studies related to models of cochlear mechanics (Sisto et al., 2007) as well as for evaluation of some factors disturbing hearing (Paglialonga et al., 2011).

One of the most advanced TF methods is matching pursuit (MP). It can identify particular frequency components of CEOAEs and provide a complete time–frequency characteristic of the signal (e.g., Jedrzejczak et al., 2008). MP gives a clear indication of where a component of a certain frequency starts and where it ends, and how its amplitude varies, making physical interpretation easier. It is possible to see how signals of defined frequency, latency, and amplitude vary with time (Jedrzejczak et al., 2009b), and certain distinctive features of CEOAEs such as brief transients (Notaro et al., 2007) and long-lasting resonances can be identified (Jedrzejczak et al., 2007). The MP method, in general, allows a detailed characteristic of latency–frequency variations to be generated and is particularly good at picking up long-lasting components known as synchronized spontaneous otoacoustic emissions, or SSOAEs (Jedrzejczak et al., 2009a).

However, turning to the linear protocol, there have been few studies using the TF method (Tognola et al., 1997; Moleti et al., 2002). There is also the series of works of Hatzopoulos and colleagues that systematically looked at TF differences between the stimulus protocols (Hatzopoulos et al., 2000a,b,c). Their main finding was a marked similarity between CEOAEs evoked by the different protocols in the 4–14 ms window (when a 70 dB SPL stimulus was used for the linear protocol and an 80 dB SPL stimulus for the nonlinear).

The purpose of the present study was to use the MP method to make precise comparisons of CEOAEs recorded under the linear and nonlinear protocols. Of particular interest were the artifacts occurring in the early time-window of CEOAEs recorded in the linear protocol.

II. MATERIALS AND METHODS

CEOAEs were recorded from both ears of 25 normal-hearing subjects using the ILO96 system (Otodynamics Ltd., Hatfield, U.K.). All subjects were female (age, 22–45). They were laryngologically healthy and had no otoscopic ear abnormalities. Impedance audiometry gave normal type-A tympanograms and normal acoustic reflexes. Hearing thresholds were better than 20 dB HL for all test frequencies (0.25, 0.5, 1, 2, 3, 4, 6, 8 kHz). Stimulus levels were kept at standard default settings of 78–82 dB pSPL. Responses were recorded using the linear protocol and the nonlinear protocol and consisted of 512 data points, stored in each of two buffers, representing 520 averages (double the standard 260 to obtain high-quality waveforms). In the nonlinear protocol, the clicks appeared in groups of four, three of the same amplitude and a fourth at three times greater amplitude and inverted polarity. Both nonlinear and linear recordings used a recording window of 2.5–20 ms. All responses (including single trials) were recorded for off-line analysis.

Additionally, each subject was tested for the presence of synchronized spontaneous OAEs (SSOAEs) using the in-built technique provided by the ILO96 equipment. An ear was classified as “with SSOAEs” when at least one long-lasting peak was found in the SSOAE spectrum that exceeded the noise floor by 5 dB SPL. There were 27 ears with SSOAEs and 23 ears without. To help isolate stimulus artifacts, measurements were also made on two deaf subjects (4 ears), also female, who had cochlear implants but intact tympanic membranes. The ear canals of the deaf subjects were presumed to have similar acoustic properties to those of the normal hearing subjects.

The Wilcoxon rank sum test was used for statistical analysis and the criterion of significance was set at $P < 0.05$. This test is largely equivalent to Student’s $t$-test and is used when the population under study does not follow a normal distribution.

A. Time–frequency (TF) analysis by matching pursuit (MP)

Time–frequency (TF) analysis of the recorded signals was done by decomposing them into their basic waveforms. For this, a method of high-resolution adaptive approximation was used, a technique based on the matching pursuit (MP) algorithm (introduced by Mallat and Zhang, 1993). This method has been shown to provide interesting insights into many physiological signals (Salamalekis et al., 2006; Sieluzycki et al., 2009), including OAEs (Jedrzejczak et al., 2004; Notaro et al., 2007). MP can decompose multi-
frequency CEOAE signals into pure-tone waveforms that can be defined in terms of their frequency, amplitude, latency, and time span. In this paper, MP with an “enriched dictionary” of symmetric and asymmetric waveforms (Jedrzejczak et al., 2009a) was used, a method developed especially for CEOAE analysis. Compared with the classical MP (which uses a symmetric Gabor dictionary) and the related wavelet transform, the main advantage of this new MP approach is that it does not generate false latency shifts in ears that have spontaneous OAEs (SOAEs, Jedrzejczak et al., 2007). The TF evolutions of the signal are herein presented in the form of amplitude distributions (Jedrzejczak et al., 2009a). Applied to OAEs, it generates time–frequency plots of high resolution in which amplitude variations of the signal are shown in terms of shading density (see for example Fig. 1). Typically, the plots contain a number of narrow and repeatable horizontal bands that are characteristic of each individual ear and which reflect stable “resonance modes” (Jedrzejczak et al., 2009a).

III. RESULTS

Examples of linear and nonlinear recordings from the same ear are shown in Fig. 1. The original time–amplitude responses are shown at the top and below them the corresponding TF distributions. At far right, the SSOAE spectrum is shown. The peak at around 5 kHz reveals a long-lasting component that corresponds with the continuous trace of energy that spans the whole width of the window on both TF plots. Importantly, the main signal components are similar for both the linear (left) and nonlinear (right) protocols but of particular interest are the differences. The main difference is the strong vertical band that shows up in the linear response at 4–5 ms (left) but that does not appear at all in the nonlinear response (right). This feature represents a component of high energy that has a latency of around 4.5 ms and contains frequencies of 0.5–2.5 kHz. In almost all our data-sets, a similar difference between the two types of maps was observed. This difference is brought out clearly in Fig. 2, which shows the TF average across all subjects. Here also the greatest difference between the two protocols appears in the short-latency, low-frequency region. The main evoked response forms a triangular shape in the nonlinear analysis (bottom) and an additional dark patch is present below the triangle in the linear analysis (top). Long horizontal lines of activity represent SSOAEs.

To find the origin of this low-frequency feature, two approaches were undertaken. The first was to make recordings from two totally deaf subjects (4 ears that all had cochlear implants). The recordings were made in the same way as for the normally hearing subjects and were done using both linear and nonlinear protocols. Because the cochleas of these deaf subjects were completely inactive, they did not generate OAEs, and the recorded signals were composed only of acoustic artifacts and noise. When the nonlinear recordings were analyzed with the TF technique, they contained no distinctive artifacts only uniformly spread noise (data not shown). In the case of the linear recordings, however, TF analysis gave, as a group average, the result shown in Fig. 3(A). The plot shows the linear average over the standard ILO recording window of 2.5–20 ms, and it can be seen that the stimulus artifact occupies much of the frequency range and lasts several milliseconds. For comparison, the insert in the top right of Fig. 3(A) shows the TF map of the same data without the standard ILO blanking window of 0–2.5 ms. Inspection of the main figure shows that signal

FIG. 1. CEOAEs recorded with both the linear protocol (left) and nonlinear protocol (right) for one subject. Top: CEOAE waveforms showing amplitude as a function of time. Bottom: TF representations (with energy shown as levels of gray) after applying the MP method. Far right: SSOAE spectrum showing long-lasting components.
energy in the linear protocol is concentrated in the range 2.5–5 ms in time and 0–2 kHz in frequency. As shown by the inset, without the 0–2.5 ms window, this artifact energy extends as far as 6 kHz.

The second approach to pinpointing the origin of the artifacts was to subtract the nonlinear TF response from the linear one on a subject-by-subject basis. Effectively, the process is equivalent to subtraction of the map of Fig. 2(B) from the map of Fig. 2(A). In other words, the nonlinear recording was treated as the reference signal and was assumed to contain no artifacts; it was subtracted from a (normalized) linear signal and the difference normalized again, a process that is somewhat arbitrary but quite robust. Again the greatest difference between the two protocols was located in the short-latency, low-frequency region. To specify this region more precisely, boundaries were drawn where the amplitude of the difference between linear [Fig. 2(A)] and nonlinear [Fig. 2(B)] responses exceeded 50% of the maximum (this difference was significant at the level of \( P < 0.01 \)). This area, shown in Fig. 3(B), spans 2.5–5 ms in time and 0–2.2 kHz in frequency and is broadly consistent with the tests on the deaf subjects [Fig. 3(A)], where a similar region is seen to contain linear artifacts. The same analysis was also performed at lower stimulus levels: The artifact remained for stimuli of 70 and even 60 dB SPL and covered a similar TF space (data not shown), consistent with reports cited earlier (Ravazzani et al., 1996; Hoth et al., 2007).

The advantage of TF analysis is that once a contributing region is identified, an MP algorithm can be brought to bear to algebraically subtract it from the original signal. That is, any dark band, representing a waveform with latencies of 2.5–5 ms and frequencies of 0–2.2 kHz, can be removed from the signal. The drawback of the procedure is that it might not be easily done without access to sophisticated signal-processing software. Fortunately, however, there is an equivalent method based on three simple steps. The first step is to filter the linear CEOAE signal below 2.2 kHz; we have found that a second-order Butterworth filter is adequate. The second step is to apply windowing from 2.5 to 5 ms using a cosine window. Finally, this derived signal can be subtracted from the linear CEOAE recording.

An example of using both procedures is shown in Fig. 4. Figure 4(A) shows the original linear OAE signal, which is windowed 2.5 to 20 ms by the ILO system. Plots in Figs. 4(B) and 4(C) show the steps involved in the MP procedure. First, Fig. 4(B) shows the artifact part of the signal extracted using MP decomposition by selecting waveforms in the area from 0 to 2.2 kHz and 0 to 5 ms. Figure 4(C) shows how MP can
then be used to reconstruct the signal after excluding these waveforms. For comparison, plots in Figs. 4(D) and 4(E) show the steps involved in the windowing and filtering method. Figure 4(D) shows the windowed and filtered signal from 2.5 to 5 ms and 0 to 2.2 kHz, and Fig. 4(E) shows the result of subtracting signal D from signal A. One can see that the artifact identified by both approaches [Figs. 4(B) and 4(D)] is very similar. Confirming the effectiveness of these techniques, the dashed line in Figs. 4(C) and 4(E) shows the nonlinear OAE signal. The two traces in each of Figs. 4(C) and 4(E) are so close, it is difficult to see the difference. The linear OAE is now free of artifact but, satisfyingly, retains its high-frequency response. The same level of performance was possible for data from all our subjects.

As emphasized in our previous work (Jedrzejczak et al., 2009a,b), MP provides a parametric description of signal components. It provides a way of readily identifying distinctive signal features, most notably frequency, amplitude, and latency. In the present study, analysis for amplitude and latency could reveal if physically significant parts of the response had been lost by processing. Figure 5 displays the amplitude and latency of MP-derived signals; they have been averaged over all subjects and appear in half-octave frequency bands between 707 and 4000 Hz.

It can immediately be seen [Fig. 5(A)] that the highest amplitudes came from the linear signal, and statistically all the data points with center frequencies from 707 to 2000 Hz had amplitudes larger than their nonlinear equivalent ($P < 0.05$). Of course, amplitude by itself is not necessarily a good indicator; more important is the amplitude of the filtered signal in relation to the nonlinear signal, the latter being the “gold standard” so far as OAEs are concerned. In this regard, the filtered signal was nearly always equal to or

FIG. 4. How stimulus artifact can be removed from an OAE signal (for one representative ear). MP method is shown in (B) and (C), and combined windowing and filtering in (D) and (E). (A) Original linear signal as recorded, windowed, and filtered by ILO system. (B) Artifactual component of OAE signal extracted by choosing waveforms from 0 to 2.2 kHz and 0 to 5 ms region of MP decomposition. (C) Reconstruction of OAE signal after exclusion of waveforms from 0 to 2.2 kHz and 0 to 5 ms (solid line) superimposed on a trace from a nonlinear recording (dashed line). (D) Artifact extracted by filtering and windowing the signal in (A) from 0 to 2.2 kHz and 0 to 5 ms (solid line). (E) (A) minus (D) (solid line) superimposed on a trace from a nonlinear recording (dashed line). The traces in (C) and (E) are so similar that the dashed line is barely visible.

FIG. 5. CEOAE amplitude (top) and latency (bottom) averaged over all subjects as determined by MP. Statistically significant differences between linear data (circles) and nonlinear data (squares) were found for amplitude and latency parameters in all frequency bands from 750 to 2828 Hz. Triangles show linear data after removal of components with $t < 5$ ms and $f < 2.2$ kHz. Bars show standard errors.
somewhat larger than the nonlinear signal, and this is generally regarded as advantageous.

When the latencies were compared [Fig. 5(B)], values in the range 0.7–2 kHz were shorter for the linear protocol compared to the gold standard ($P < 0.02$). But after the components connected with the stimulus artifact were removed (according to the procedure described in the preceding text), the differences between the filtered signal and the nonlinear data were very small and not statistically significant. In fact, in both cases, the latency–frequency plots form an approximate straight line (on logarithmic axes), fitting the linear power law proposed on theoretical grounds by earlier workers (Tognola et al., 1997).

The main motivation for using the linear protocol is because it is expected to produce more robust responses (Hatzopoulos et al., 2003). But is this still true after removal of artifacts? To answer this question, standard parameters such as reproducibility and signal-to-noise ratio (SNR) were evaluated, and the results are plotted in Fig. 6. An additional part of this analysis was to compare ears with SSOAEs (27 ears) and those without (23 ears). Reproducibility was calculated as the correlation between the sub-averages stored in the two acquisition buffers (Kemp et al., 1986) and were expressed in percent ([Figs. 6(A) and 6(C)]. SNR was evaluated as the difference in dB between responses and the noise [Figs. 6(B) and 6(D)]. Of course, at a basic level reproducibility and SNR are related (Tognola et al., 2001b; Korres et al., 2006). The plots of Fig. 6 show that after removal of the stimulus artifact by filtering and windowing, both SNR and reproducibility of the linear protocol remained significantly higher ($P < 0.01$) than for the reference nonlinear protocol. These differences were at least 6 dB and 5%, respectively, but for some frequency bands (especially the higher ones) they reached as much as 12 dB and 10%.

In addition, Fig. 6 shows a further clear-cut difference between the linear and nonlinear protocols: the nonlinear protocol appears to be systematically affected by the presence of SSOAEs. In nearly all frequency bands, CEOAE responses from ears with SSOAEs (black bars) had higher reproducibility and SNR ($P < 0.05$) than ears without

![Graphs A, B, C, D showing reproducibility and SNR comparisons](https://example.com/graphs.png)
SSOAEs (white bars). Although the linear protocol also tended to show the same disparity, the difference was small and not statistically significant. It is also worth noting that the figures presented here derive from an average of 520 responses; comparable figures for the standard 260 average in the ILO system would be about 5% and 5 dB lower.

A further clue as to what might be going on here is found in Fig. 7, which pools all the frequency bands and shows how the global reproducibility for linear and nonlinear data increases as a function of the number of averaged trials. Again, ears with and without SSOAEs are distinguished. (In order to make the data from linear and nonlinear data comparable, each subaverage is treated as the mean of four responses: that is, for the nonlinear protocol, the subaverage represents responses evoked by three stimuli of the same level and polarity and the fourth of three times greater level and reversed polarity; for the linear protocol, however, the subaverage derives from four stimuli of the same amplitude and polarity.) As Fig. 7(A) makes clear, nonlinear signals from ears with SSOAE show significantly higher global reproducibility ($P < 0.05$) than ears without SSOAEs. It is notable that for ears without SSOAEs, the 90% criterion of reproducibility cannot be achieved even after 520 subaverages, double the default number of the ILO system. By way of contrast, the results of the linear method, with additional windowing and filtering, are shown in Fig. 7(B). In this case, the difference between ears with and without SSOAEs was smaller and not significant, but the important thing is that the 90% reproducibility criterion was reached after less than 100 averages.

IV. DISCUSSION

We have recorded CEOAE waveforms under both the linear and nonlinear protocols and have used the MP method to achieve high TF resolution over a signal space of 2.5–20 ms and 0–6 kHz. Our main finding is that when using stimuli of around 80 dB SPL, the clearest difference between the two protocols can be traced to the 2.5–5 ms window and is probably due to stimulus artifacts, a result consistent with Ravazzani et al. (1996). This interpretation is confirmed by our comparison of recordings made from healthy ears and from ears without OAEs.

A noteworthy result is that the stimulus artifact is limited in frequency to about 0–2.2 kHz (Fig. 3). When we compare the area contaminated by artifact with the results of TF analysis, we find that signals approaching the form of those derived by the standard nonlinear protocol can be recovered. That is, from Fig. 4 we see that OAE responses above 2.2 kHz can be recovered from the 2.5–5 ms window by simple high-pass filtering, while OAEs below 2.2 kHz can be recovered from the window beyond 5 ms. Our conclusion is that it is possible to remove stimulus artifact without losing the high-frequency OAEs. There are two ways of removing this artifact. It can be done using a TF method, such as MP, but an easier way is through windowing and linear filtering. Both these methods can be applied to existing recordings. Because of its simplicity, the windowing and linear filtering method could easily be incorporated into the data acquisition stages of OAE devices. It is also worth noting that artifacts can sometimes also be present in nonlinear recordings (e.g., Whitehead et al., 1995), and the same procedure could be used to cancel them out.

When the stimulus artifact is removed in the way we propose, the signal amplitude is reduced to levels similar to those in the nonlinear protocol. The TF properties also become similar (Fig. 5). This suggests that a large portion of the stimulus artifact was indeed removed and that the processing did not change the overall properties of the OAE signal. At the same time, reproducibility and SNR were still higher than for the nonlinear protocol (Fig. 6). Whether the processed signal is more diagnostically useful than the nonlinear signal will need further study involving populations of subjects with known hearing deficits.

An important foundation of our method is the assumption that the artifact has invariable time–frequency properties, and this aspect requires some justification. A known characteristic of CEOAEs is that high frequency components have short latencies and low frequency components have longer...
ones. This latency–frequency dependence can be described by a power law of the form \( t = a f^{-b} \) (Tognola et al., 1997), where \( t \) is the latency, \( f \) is the frequency, and \( a \) and \( b \) are constants. Experimentally, \( a \approx 9.2 \) and \( b \approx 0.5 \), yielding delays of around 9.2 ms at 1 kHz, 6.5 ms at 2 kHz, and 4.6 ms at 4 kHz, values which this study generally confirms. These latency figures have been determined for both linear (Tognola et al., 1997) and nonlinear (Jedrzejczak et al., 2009a) protocols; however, it should be pointed out that there have been no studies that have compared latency and frequency on the same set of subjects using different protocols.

Nevertheless, turning to the data presented here, it is clear that before artifact removal, the latencies under the linear protocol are generally smaller than those under the nonlinear protocol (Fig. 5). Moreover, the latency at 0.7 kHz is always smaller than at 1 kHz, a clear anomaly in terms of the power law. This anomaly can also be seen in other TF studies of CEOAEs performed under the linear protocol (Tognola et al., 1997). Significantly, our Fig. 5 shows that after the artifact is removed, the latency–frequency relationship for the linear protocol becomes similar to that of the nonlinear protocol and the anomaly disappears.

Another aspect worth noting is that our study was done using the widely used ILO system, and different equipment is likely to generate different artifacts, particularly if it does not use the standard 0–2.5 ms blanking window. Although our Fig. 3(A) shows that all frequencies in the 0–2.5 ms window are contaminated by artifact, other set-ups will probably generate their own sort of artifacts and require different removal methods. For example, the system used by Goodman et al. (2009) was able to record very short latencies and was able to extract OAEs up to 16 kHz. This work pointed out that the early component may be of the generator type while later ones may be of the reflection type. The generator component has greater amplitude and grows linearly with stimulus level compared to the reflection component, which grows more compressively. The present study did not seek to separate the two sources, but we acknowledge that this factor could mean that windowing of 0–5 ms and filtering of 0–2.2 kHz may favor one component over another, and this possibility needs investigation. Nevertheless, Fig. 5 supports the idea that windowing and filtering makes the linear response contain signal sources similar to those measured with the nonlinear protocol.

One outstanding disadvantage of all types of OAEs is their variability between subjects. In the case of distortion product OAEs, it is possibly due to the interaction of generator and reflection components (e.g., Whitehead et al., 1992; Long et al., 2008). In the case of CEOAEs, our work suggests (Fig. 6) that the evoked part of the response appears to be mixed with SSOAEs, an idea that fits in with the finding that ears without SOAEs tend to have lower levels of CEOAEs (Kulawiec and Orlando, 1995; Prieve et al., 1997; Schmuziger et al., 2005). This explains our finding that SSOAEs affect reproducibility and SNR, particularly for the nonlinear protocol. Previous workers have attempted to improve reproducibility by separating the evoked and spontaneous components of CEOAEs (Bennett and Ozdamar, 2010); other possible approaches have involved improving stimulation and acquisition protocols (Yang et al., 2003; Goodman et al., 2009) or using advanced signal processing techniques (Jedrzejczak et al., 2007). In the light of our findings, a better alternative may be to use the linear protocol, which appears to be only slightly affected by SSOAEs. Although SSOAEs are present in OAEs recorded under both the linear and nonlinear protocols (Fig. 1), we find that the linear protocol provides OAEs of significantly less variability (Fig. 6) with or without SSOAEs.

Finally, there is a practical matter. The quality of OAE recordings is strongly affected by the number of averaged responses (because SNR is proportional to its square root). A prime advantage of linear-based method is that it improves SNR by at least 6 dB in each half-octave frequency band compared to the nonlinear method (Fig. 7). In practical terms, this means that the number of averaged CEOAE responses needed to reach a “pass” in the ILO system will be less. This is an aspect not often discussed in the literature, but it is important during testing of newborns when a short test-time is preferred (Korres et al., 2006). In the case of the ILO equipment, the standard number of test responses is 260 (Prieve et al., 1993), but in screening mode, the averaging is stopped automatically after a “pass” level of 75% reproducibility is reached. Use of the linear protocol means this level can be achieved appreciably faster.

The present report should be treated as a pilot study. It shows the potential of the method on cooperative subjects under controlled conditions. Whether it can be used on newborns in a clinical setting is yet to be determined. Adults and newborns have different OAEs in terms of spectral content and prevalence of SOAEs (e.g., Keefe and Abdala, 2011), and the smaller ear canal of newborns may generate different artifacts.

V. CONCLUSIONS

Our study of CEOAEs involved a comparison of the linear and nonlinear protocols of data acquisition. The main findings can be summarized as follows.

(1) TF analysis has shown that the main difference between CEOAEs recorded with the linear and nonlinear methods is found in the 2.5–5 ms post-stimulus time window, and it affects frequencies below 2.2 kHz. The difference is mainly due to stimulus artifact in the linear protocol.

(2) This artifact can be removed, either by the MP method or by a combination of windowing and linear filtering, a process that still leaves the high-frequency, early OAE response intact. TF analysis verifies the reliability of the proposed method.

(3) After applying our method of artifact removal, the TF properties of CEOAEs recorded under both the linear and nonlinear protocols were similar. Reproducibility and SNR were still higher for the linear CEOAEs.

(4) CEOAEs recorded in the linear protocol were less influenced by the presence of SSOAEs.

(5) Our proposed method for removing the stimulus artifact may improve the reliability of CEOAE measurements in the linear protocol when an 80 dB SPL stimulus is used.
