Hearing Research 335 (2016) 161–178

Contents lists available at ScienceDirect

Hearing Research

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

Research paper

Time–frequency decomposition of click evoked otoacoustic emissions in children

Srikanta K. Mishra [a], [*], Milan Biswal [b]

[a] *Department of Special Education and Communication Disorders, New Mexico State University, Las Cruces, USA*
[b] *Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, USA*

ARTICLE INFO

Article history:
Received 8 December 2015
Received in revised form
25 February 2016
Accepted 7 March 2016
Available online 11 March 2016

Keywords:
Click evoked otoacoustic emissions
CEOAE latency
CEOAE level
S transform
Time–frequency distribution
Wavelet transform

ABSTRACT

Determining the time–frequency distributions of click-evoked otoacoustic emissions (CEOAEs) are scientifically and clinically relevant because of their relationship with cochlear mechanisms. This study investigated the time–frequency properties of CEOAEs in 5–10 year old children. In the first part, we examined the feasibility of the *S* transform to characterize the time–frequency features of CEOAEs. A synthetic signal with known gammatones was analyzed using the *S* transform, as well as a wavelet transform with the basis function used traditionally for CEOAE analysis. The *S* and wavelet transforms provided similar representations of the gammatones of the synthetic signal in the mid and high frequencies. However, the *S* transform yielded a slightly more precise time–frequency representation at low frequencies (500 and 707 Hz). In the second part, we applied the *S* transform to compare the time–frequency distribution of CEOAEs between adults and children. Several confounding variables, such as spontaneous emissions and potential efferent effects from the use of higher click rates, were considered for obtaining reliable CEOAE recordings. The results revealed that the emission level, level versus frequency plot, latency, and latency versus frequency plot in 5–10 year old children are adult-like. The time–frequency characteristics of CEOAEs in 5–10 year old children are consistent with the maturation of various aspects of cochlear mechanics, including the basal to apical transition. In sum, the description of the time–frequency features in children and the use of the *S* transform to decompose CEOAEs, are novel aspects of this study. The *S* transform can be used as an alternative approach to characterize the time–frequency distribution of CEOAEs.

1. Introduction

Click evoked otoacoustic emissions (CEOAEs) are widely used in newborn hearing screening programs and pediatric auditory diagnostics. However, their basic temporal features, such as delay or latency, are not thoroughly characterized in children. CEOAEs are considered a type of reflection source emission (Shera and Guinan, 1999). The coherent reflection theory postulates that reflection emissions arise from the backscattering of energy off hypothesized cochlear micromechanical inhomogeneities (Shera and Guinan, 2003, 1999; Shera et al., 2008; but see Ren, 2004; Siegel et al., 2005). However, at higher stimulus levels, reflection emissions may receive contributions from the basal cochlear regions of coherent reflection (Lewis and Goodman, 2015; Sisto and Moleti, 2008; Withnell et al., 2008). Since reflection source emissions are generated near the peak of the traveling wave, OAEs recorded at low levels are sensitive to subtle cochlear amplifier gain changes (Shera, 2004). For example, medial efferent activation by contralateral acoustic stimulation produces larger changes in the reflection-compared to the distortion-component of distortion product OAEs (Abdala et al., 2009; Mishra and Abdala, 2015). Additionally, reflection-source OAE delays, including CEOAE latencies, can be applied to objectively estimate cochlear tuning in humans and other animals (Bentsen et al., 2011; Bergevin et al., 2012; Joris et al., 2011; Keefe, 2012; Moleti et al., 2008; Shera et al., 2002; Sisto and Moleti, 2007). The latency at a given frequency is defined as the time interval between the onset of the stimulus and the frequency-specific component of CEOAEs. The

---

Abbreviations: CEOAEs, click evoked otoacoustic emissions; dB, decibel; HL, hearing level; MSE, mean squared error; OAEs, otoacoustic emissions; *pe*SPL, Peak-equivalent sound pressure level; SNR, signal-to-noise ratio; SOAEs, spontaneous otoacoustic emissions; ST, *S* transform; WT, wavelet transform

[*] Corresponding author. Department of Special Education and Communication Disorders, New Mexico State University, Las Cruces, NM 88003, USA.

E-mail address: smishra@nmsu.edu (S.K. Mishra).

http://dx.doi.org/10.1016/j.heares.2016.03.003

latency is difficult to compute in CEOAEs due to the presence of numerous frequency components in the emissions. However, decomposing the original CEOAE response into a set of spectral components allows for the computation of latencies at specific frequencies (e.g., Tognola et al., 1997).

The description of the basic features of CEOAEs in children is nearly two decades old (e.g., Bray and Kemp, 1987; Norton and Widen, 1990; Prieve et al., 1997). Previous studies suggest that CEOAE levels measured in children are higher than those measured in adults. Bray and Kemp (1987) argued that the higher OAE levels in 6–13 years old children, relative to adults, could be attributed to the differences in their ear canal volumes. Norton and Widen (1990) reported significant differences in CEOAE levels evoked by 80 dB peak SPL clicks among the studied age groups; 0.0–9.9, 10.0–19.9, and 20.0–29.9 years. Similarly, Prieve et al. (1997) showed that children aged 1–5 years had higher CEOAE levels than those aged 12–17 years and adults. These early studies provide a good description of CEOAE levels, but are mostly recorded at high stimulus levels (but see, Prieve et al., 1997) and with a non-linear recording mode (Kemp et al., 1986). However, the time–frequency properties of CEOAEs, including the latency or delay, remain to be adequately characterized in children. Limited data suggests that CEOAE latencies, recorded at high click levels using a non-linear recording method, are longer in newborns compared to adults (Moleti and Sisto, 2003; Moleti et al., 2008). However, it is unknown whether similar trends can be observed for children representing other age groups, and for a range of recording parameters.

Characterizing the CEOAE latency in children is not only important from a developmental perspective, but also has significant translational implications. For instance, CEOAE latency can be used to objectively estimate cochlear tuning in the pediatric population, in whom psychophysical tuning curves may be difficult to obtain (Moleti and Sisto, 2003; Moleti et al., 2008). The latency can be applied to accurately characterize the efferent effects on cochlear mechanisms (Francis and Guinan, 2010) and compute a vector metric to index efferent reflex (Abdala et al., 2013; Marshall et al., 2014; Mishra and Abdala, 2015).

From a signal processing perspective, CEOAEs are non-stationary signals that exhibit frequency dispersion, meaning that different frequencies propagate at different phase speeds. The low frequency components have longer delays compared to high frequency components, consistent with the place-frequency map of the cochlea (Kemp, 1978). Additionally, according to the coherent reflection theory, the frequency content of the signal changes with time (Shera and Guinan, 1999; Zweig and Shera, 1995). Time–frequency decomposition of OAE signals provide information simultaneously in both the time and frequency domains, which may be useful for their interpretation, but which may not be available in spectral or temporal analyses. For example, the fast Fourier transform does not provide information on frequency changes along time, and fails to represent the fine changes in the CEOAE waveform that are direct consequences of the time-varying behavior of the signal. In contrast, the time–frequency analyses show the energy distribution of the evoked emissions as a function of both time and frequency, facilitating the computation of the latency of individual spectral components. Therefore, accurately characterizing the time–frequency distribution of CEOAEs may provide important insights into cochlear mechanisms, such as the generation mechanism of OAEs (e.g., Sisto et al., 2015) and cochlear frequency selectivity (e.g., Moleti and Sisto, 2003).

A basic approach for analyzing non-stationary signals is the short-time Fourier transform. This technique captures the changes in frequency over time by using a window function with a fixed width to provide temporal representation. However, this method faces problems related to the choice of width of the window function (i.e., longer windows provide better frequency resolution but poorer time resolution and vice versa if shorter windows are used). In CEOAE literature, several time–frequency approaches, such as the short-time Fourier transform (e.g., Francis and Guinan, 2010), matching pursuit (e.g., Jedrzejczak et al., 2009), wavelet transform (WT) (e.g., Narne et al., 2014) and Wigner distribution (e.g., Konrad-Martin and Keefe, 2003) were applied to describe time–frequency representation. However, with the exception of the WT (Tognola et al., 1998), performance comparisons among different time–frequency transform approaches were not adequately addressed in analyzing CEOAEs. Performance of a particular time frequency analysis approach, among other things, depends on the nature of the signal (Wacker and Witte, 2013). Therefore, an approach suitable for analyzing electroencephalogram signals may not give the desired results for OAE signals.

One of the approaches for describing the time–frequency distribution of signals is the *S* transform (ST) (Stockwell et al., 1996). The ST is a linear time–frequency analysis technique which has found extensive applications for analyzing non-stationary signals in multiple disciplines ranging from geological signals, to power signals to biomedical signals (Assous and Boashash, 2012; Biswal and Dash, 2013a, 2013b; Pinnegar and Mansinha, 2003; Pinnegar et al., 2009; Raković et al., 2006; Wu et al., 2010). The kernel of the ST approach is defined by a combination of the Fourier basis and a Gaussian window. The ST can be interpreted as a special case of a continuous WT with additional advantages, such as absolute phase reference and fast computations (Ventosa et al., 2008). The absolute phase reference is a unique feature of the ST, which enables the extraction of accurate phase information relative to a fixed time reference or to the starting point of the signal. In contrast, phase information extracted from the WT is relative to the local analyzing wavelet. The computational complexity of the ST is high (Brown et al., 2005). In order to address the issue related to the computational demands, a low complexity framework for a fast computation of the ST has been proposed to analyze electroencephalogram signals (Brown and Frayne, 2008; Brown et al., 2010). Additional fast, improved and generalized versions of the ST have been recently suggested and applied to analyze non-stationary signals (Biswal and Dash, 2013a, 2013b). In this paper, we evaluated the utility of a fast variant of the ST (Biswal and Dash, 2013a, 2013b) to extract the time–frequency distribution of CEOAEs.

The two necessary requirements that a time–frequency approach must satisfy for characterizing the features of CEOAE signals are: (a) an accurate estimation of the constituent, time varying spectral components, i.e., accurate measurement of duration and bandwidth of gammatones, and (b) to obtain a high energy concentration in the time–frequency distribution with minimal spectral leakage effects (Tognola et al., 1998, 1997). In this paper, we focus on these requirements to present a comparative performance of the ST and continuous WT, strictly from the perspective of a CEOAE analysis with a simulated signal, based on the gammatone model. Because the exact, constituent spectral components of real CEOAE signals are unknown, a comparative evaluation of time–frequency approaches with real CEOAEs is not possible. Therefore, we used a synthetic CEOAE model with known spectral and temporal characteristics, and applied it as a test signal to compare the estimation accuracies of the ST and continuous WT. The WT was chosen for comparison because this method has been extensively used in CEOAE literature (Bhagat et al., 2013; Narne et al., 2014; Tognola et al., 1997).

The overall goal of this study was to characterize the time–frequency distribution of CEOAEs in children. The specific aims were (1) to demonstrate the feasibility of the ST for analyzing CEOAEs, by comparing the ST and the WT (Tognola et al., 1998,