Technical University of Denmark

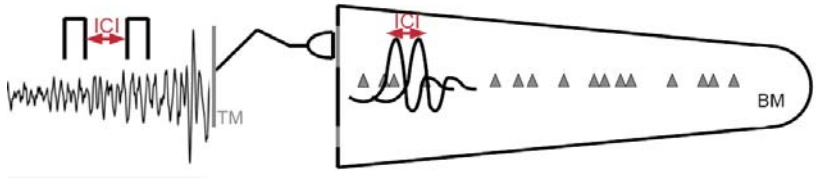
DTU

CONTRIBUTIONS TO
HEARING RESEARCH

Volume 8

Sarah Verhulst

Characterizing and modeling dynamic processes in the cochlea using otoacoustic emissions



Centre for Applied Hearing Research
Department of Electrical Engineering

Characterizing and modeling dynamic processes in the cochlea using otoacoustic emissions

PhD thesis by

Sarah Verhulst

Technical University of Denmark
2010

ISBN 978-87-92465-24-5
Printed in Denmark by Rosendahls - Schultz Grafisk a/s

Preface

This thesis was submitted to the Technical University of Denmark (DTU) as partial fulfillment of the requirements for the degree of Doctor of Philosophy (PhD) in Electronics and Communication. The work presented in this thesis was completed between December 1st, 2006 and April 15th, 2010 at the Centre for Applied Hearing Research (CAHR), Department of Electrical Engineering, DTU. An external research stay at the Eaton-Peabody Laboratory of Auditory Physiology, Harvard Medical School was undertaken in spring 2009 as a part of this PhD project. The work presented in this thesis was supervised by Prof. Torsten Dau and Assoc. Prof. James Harte at DTU, and by Assoc. Prof. Christopher Shera at the Eaton-Peabody Laboratory. A public defense was held on June 10^{th} 2010 with Prof. David Kemp, Prof. Stephen Neely and Assoc. Prof. Jörg Buchholz, as members of the assessment committee. The project was jointly funded by GN ReSound A/S, DTU and Forskerskolen SNAK. The external research stay was furthermore supported by the Oticon and Otto Mønsted Foundation. Ethical considerations for conducting the experiments at DTU were made with respect to the Science Ethics Committee for the Capital Region of Denmark; reference H-KA-04149-g.

Writing this thesis has been an interesting learning process, that I have greatly enjoyed. I have learned so many things during this PhD project, still there is so much more to explore. Even though this thesis only gives a small contribution to the greater body of auditory research, I hope that some part of the evolution is reflected in this work. I hope you enjoy reading this thesis.

Sarah Verhulst, 15^{th} of April 2010.

Acknowledgment

The quality of this thesis has been greatly improved through the supervision of Torsten Dau, James Harte and Christopher Shera. Thank you Torsten for having me in the group, I have had a great time at CAHR! From the first lectures I received from you, I knew that this was the research drive I was after. Thanks for inspiring, asking the "why's", and for showing me what it takes to become a researcher. Thanks James for all your time and effort, and for showing me how to conduct a scientific experiment. Thank you Chris for having me over in Boston, and for helping me out so much with the paperwork and my research. It was a pleasure to discuss with you and learn from you. It has been a nice experience to work with all of you, and even though I have not finished learning, you have brought me advice and skills that will help me along the way.

I would like to thank Hendrikus Duifhuis, Peter van Hengel and Bastian Epp for sharing the Fortran code of the original cochlear model, and for getting me started with the simulations. I would also like to thank my test subjects who spent so many hours "relaxing for science" in the measurement booth. Tak Morten for at oversætte mit resumé til dansk. A special thanks also goes out to my colleagues and friends at CAHR. I have really enjoyed our seminars and discussions in the office and cafeteria. You were always there with a helping hand and good advice! Thanks to Rindy, Jeff, Ann, Harrison and the other EPL people for turning my stay in Boston into such a great experience.

I would like to say thank you to my family and friends in Belgium. Even though you may have not understood a word of what I was doing in DK all the time ;-), you were always there for me. Merci Sylvain for your endless support. You and the crew of DK friends did a great job in bringing some fun into the life of a PhD student. Thanks also to the little one growing inside of me, for not making me too sick in the mornings while I was writing this thesis.

Abstract

An important characteristic of human hearing is that it amplifies weak sounds while attenuating louder ones. This gain transformation takes place in the inner ear (i.e., cochlea), and is responsible for a compressive relation between the level of the presented and perceived sound. The cochlear gain mechanism is essential for our hearing and degrades when hearing impairment develops. A comprehensive understanding of the gain involved in the intact human cochlea is crucial, as hearing instruments try to compensate for the loss in cochlear gain caused by hearing damage. This thesis investigates dynamic, or time-dependent, properties of cochlear gain. A time scale from 0 to 10 ms is considered to ensure that cochlear processing is investigated without including influences from higher stages in the brain. The results are expected to provide insight into how e.g. onsets of sounds are processed by the intact human system.

Click-evoked otoacoustic emissions (CEOAEs) were used as a non-invasive technique to investigate cochlear gain mechanisms. CEOAEs are echoes to click stimuli that can be recorded in the ear canal, and that are produced in the cochlea as a byproduct of the nonlinear gain mechanism. Experimental results demonstrated that the CEOAE level-curve, i.e. the relation between click and CEOAE level, altered when a click was presented close in time to the evoking click. This effect was named "temporal adaptation" of the CEOAE level-curve, and was shown to operate on a time scale of 0 to 8 ms. The relation between dynamic features of CEOAEs and the underlying cochlear gain mechanism was furthermore investigated by means of a numerical model of the cochlea that simulates CEOAEs. The simulations provided insight into level-dependent features of the cochlear gain mechanism that underlie the generation of the CEOAE. In order to account for key features of temporal adaptation in CEOAEs, the existence of a time dependence in the cochlear gain mechanism was suggested. Overall, this study has demonstrated that cochlear compression characteristics can change on a time scale of 0–8 ms. The existence of such a time constant in cochlear compression may be of interest for the future development of signal processing in hearing instruments.

Resumé

En af de vigtigste egenskaber i menneskets hørelse er evnen til at forstærke svage lyde imens stærke lyde dæmpes. Denne forstærkning finder sted i det indre øre, dvs. cochlea, og forårsager en kompressiv sammenhæng mellem det afspillede og opfattede lydniveau. Forstærknings-mekanismen i cochlea er essentiel for hørelsen og forværres når et høretab opstår. Det er vigtigt at opnå en omfattende forståelse af forstærkningen i den intakte cochlea, fordi høreapparater forsøger at kompensere for tabet af forstærkning i cochlea som følge af et høretab. I denne afhandling undersøges de dynamiske, dvs. tidsafhængige, egenskaber for forstærkningen i cochlea. Der fokuseres på de første 0 til 10 ms, hvorved det kan antages at den undersøgte signalbehandling af cochlea ikke er påvirket af højere trin i hjernen. Det forventes at resultaterne belyser hvorvidt, f.eks. begyndelsen af lydsignaler bliver behandlet i det intakte høresystem.

For at undersøge cochleær forstærkningen blev der brugt såkaldte klik-fremkaldte otoakustiske emissioner (click-evoked otoacoustic emissions, CEOAE) som indgår i en ikke-invasiv måle-metode. CEOAE metoden måler ekkoer i ørekanalen som opstår ved stimulering med klik-signaler, og er et bi-produkt af den ikke-lineære forstærknings-mekanisme i cochlea. De eksperimentelle resultater viser at niveau-kurven for CEOAE signaler, dvs. sammenhængen mellem klik-niveau og CEOAE-niveau, ændrede sig når et klik blev afspillet tidsmæssigt tæt på det CEOAE-fremkaldende klik. Denne effekt blev kaldt "tidsmæssig adaptation" af CEOAE niveau-kurven, og det vidste sig at denne effekt gjorde sig gældende over en tidsudstrækning fra 0 til 8 ms. Sammenhængen mellem de dynamiske egenskaber for CEOAE og den underliggende cochleære forstærkning blev yderligere undersøgt med udgangspunkt i en numerisk model af cochlea som kan simulere CEOAE. Disse simulationer belyser de niveau-afhængige egenskaber for forstærkningsmekanismen som lægger til grund for genereringen af CEOAE signalerne. For at kunne beskrive for de vigtigste egenskaber i forbindelse med tidsmæssig adaptation af CEOAE signaler blev det foreslået at forstærknings-mekanismen er tidsafhængig. Samlet set har dette arbejde vist at egenskaberne for cochleær kompression kan ændre sig over en tidsudstrækning fra 0 til 8 ms. Eksistensen af en sådan tidskonstant kan vise sig interessant for fremtidens udvikling af signalbehandling i eksempelvis høreapparater.

Contents

List of abbreviations and symbols

BF	best frequency
BM	basilar membrane
CEOAE	click-evoked otoacoustic emission
CF	critical frequency
DC	double click
DCI	double click inverted
DPOAE	distortion-product otoacoustic emission
DS	derived suppressed (response)
IHC	inner hair cell
I/O function	input/output function
IR	impulse response
L_s	suppressor-click level
L_t	test-click level
OAE	otoacoustic emission
OHC	outer hair cell
peSPL	peak-equivalent sound pressure level
rms	root mean square
S-click	suppressor click
SFOAE	stimulus-frequency otoacoustic emission
SOAE	spontaneous otoacoustic emission
SPL	sound pressure level

SSOAE	synchronized spontaneous otoacoustic emission
T-click	test click
TEOAE	transient-evoked otoacoustic emission
US	unsuppressed (response)

1

General Introduction

When sound enters the ear-canal, its pathway leads through the middle and inner ear towards the brain where it is perceived, interpreted and acted upon. The outer and middle ear transform the incoming air pressure variation to mechanical vibrations of the middle-ear bones, that transmit the vibrations to the inner ear (cochlea). The bony and fluid-filled cochlea has the shape of a snail shell and transforms the mechanical vibrations it receives from the stapes (i.e., last middle-ear bone) to a complex vibration pattern along the basilar membrane (BM). The BM spans the cochlea from the base (i.e., at the middle-ear connection) to the apex (i.e., at the end of the coiled structure) and is surrounded by fluid. The vibration patterns that are created along the BM can be detected by sensors, i.e. the inner hair cells (IHCs) that are placed closely to the BM along the length of the cochlea. The synapses attached to the IHCs translate the vibration of the BM into electrical pulses that are gathered in the auditory nerve, and sent to the higher stages of the brain where perception is formed.

The vibration pattern on the BM is partly determined by the impedance properties of the BM and its surrounding fluids. The mechanical properties of the BM cause high-frequency sounds to excite the membrane near the base, and low-frequency sounds to excite the BM maximally near the apex (Greenwood, 1961). The BM can thus be seen as a cascade of "auditory filters", each tuned to a frequency that corresponds to a specific location on the membrane. The BM vibration pattern is furthermore influenced by an active process that constitutes the outer hair cells (OHCs), placed next to the IHCs along the length of the BM. OHCs are involved in a nonlinear gain mechanism that pumps more energy in vibration patterns to low-intensity than to high-intensity stimuli (e.g., Robles et al., 1986; Ruggero et al., 1997). OHCs thus "tune" the BM vibration pattern such that the IHCs can detect the individual frequency components in the pattern well. Hearing impairment (e.g., due to acoustic trauma)

is often accompanied by a loss of OHCs, which results in less sharply tuned BM vibration patterns (Glasberg and Moore, 1986). As a consequence, it becomes more difficult to segregate closely spaced frequency components in sounds, often leading to degraded speech intelligibility.

The role of the OHCs in the normal and impaired hearing system is crucial for the development of hearing prostheses, designed to compensate for the loss of sensitivity and BM tuning. It is challenging to investigate the function of the OHCs for human hearing, as the active gain mechanisms of the OHCs disappear post-mortem. Invasive measurements in mammals have been successful in providing valuable information about the gain properties of the OHCs, but do not describe the functioning of the OHCs in an intact human auditory system. Fortunately, Gold (1948) tackled the cochlear tuning properties from an engineering point of view and hypothesized that a healthy cochlea would emit sound. He found sharp tuning in the largely damped viscous system of the cochlea, and postulated the existence of an active feedback mechanism to counteract the damping, such that sharp tuning could be achieved. At cochlear locations where the active feedback would exceed the amount of damping to counteract, sustained oscillations would occur and travel out of the cochlea. The recording equipment at the time did not allow for a successful recording of these "echoes", such that Gold (1948) could not prove his theory. In the 70's, Kemp (1978) picked up the topic and was among the first to record and evoke otoacoustic emissions (OAEs). Their origin was traced back to the active mechanisms (i.e., OHCs) inside the cochlea, and OAEs have been studied extensively ever since (see Kemp (2007) for a review). OAEs have, for example, been used to investigate BM tuning properties (e.g., Shera et al., 2002), and level-dependent compression properties of the OHCs (e.g., Kemp, 1978; Schairer et al., 2006). OAEs thus provide a powerful research tool, complementary to the commonly adopted perceptual experiments that investigate nonlinear properties of cochlear processing. These nonlinear properties reflect, for example, the relation between the level of a pure-tone masker and the level of a detection tone that is just audible when the masker is presented. (e.g., Moore et al., 1984).

Otoacoustic emissions were used in this project to investigate dynamic, or time-dependent, features of human cochlear nonlinearities. An example of such a time-

dependent process can be found in fast OHC adaptation, that describes a time constant (0–20 ms) related to OHC motility (i.e., the change in OHC length when the voltage across its membrane is changed; Brownell et al., 1985). When an OHC is deflected, it takes an adaptation period before a steady state mechanical-to-electrical transduction of the OHC is obtained (Hudspeth and Gillespie, 1994). Within this time period, the OHC may thus show different compression characteristics. The existence of a time constant in the mechanical-to-electrical transduction process of OHCs suggests that the gain applied to the BM vibration pattern could change over time. To date, it is unknown whether a healthy human cochlea reveals such short time-constants in its active gain mechanisms, and how these would influence our perception. An improved understanding of the dynamic properties of cochlear nonlinearities would not only lead to an improved understanding of how stimulus onsets are processed, it could also have implications for the design of hearing instruments. Gain-application strategies in hearing instruments often include time constants such that gain is not applied instantaneously when a stimulus is detected. The implementation of such time constants has been shown to improve speech intelligibility (e.g., Moore et al., 2004), but the exact choice of constants is often only based on sound quality judgments. An improved knowledge of the time constants involved in active cochlear mechanisms may be helpful for the future development of hearing prostheses.

Otoacoustic emissions are well suited to investigate dynamics in human cochlear nonlinearities, as their origin is directly linked to the gain mechanism in the OHCs. If OHC dynamics can alter the BM vibration pattern to an incoming sound, then these dynamic changes should also be observed in the compression properties of the reflected OAEs. In the 80's, Kemp and Chum (1980) presented emission results that pointed to a short time-constant (0–10 ms) in the active mechanisms underlying the OAE. The amplitude of a click-evoked otoacoustic emission (CEOAE) was shown to be reduced or "suppressed" when a second click was presented close in time to the evoking click (Kemp and Chum, 1980). The initial experiment has been repeated and extended several times (Tavartkiladze et al., 1994; Lina-Granade and Collet, 1995; Kevanishvili et al., 1996; Kapadia and Lutman, 2000b; Hine and Thornton, 2002), but it remains unclear whether dynamic changes in cochlear nonlinearities can account for the amplitude alterations observed in the CEOAE. The purpose of this study was to

investigate the cochlear mechanisms that underly the temporal suppression experiment originally conducted by Kemp and Chum (1980). As Kemp and Chum's experiment was the first (and one of the few) to investigate cochlear dynamics on a time scale of 0–10 ms, it is expected that the outcome of this project will lead to an improved understanding of the dynamics involved in human cochlear processing.

Chapter 2 focuses on the methodology behind the temporal suppression experiment. Two of the previous studies (Tavartkiladze et al., 1994; Hine and Thornton, 2002) occasionally reported increased levels of the CEOAE when presenting a click 0–10 ms before the test click , whereas the other studies (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kevanishvili et al., 1996; Kapadia and Lutman, 2000b) only reported level reductions. As these conflicting results may lead to different interpretations of the underlying cochlear mechanisms, chapter 2 explores and improves the different recording techniques that have been used in the past. A phenomenological model that is able to account for augmentation (i.e., increased CEOAE levels) as well as suppression (i.e., decreased CEOAE levels) is furthermore proposed and discussed.

Chapter 3 uses an optimized measurement method to investigate temporal suppression. Much like the level-dependent compression observed in BM vibration patterns recorded in animals (e.g., Robles et al., 1986), CEOAE levels show increased compression for increased stimulus levels (Kemp and Chum, 1980). It is hypothesized that the CEOAE level-curve, reflecting the relation between stimulus and CEOAE level, should show a dynamic behavior if two clicks, separated in time, change the cochlear nonlinearities that underly the generation of the CEOAE. The results from this chapter are expected to lead to an improved description of the dynamics in CEOAEs. The results are furthermore used to demonstrate that dynamic changes in the CEOAE level-curve (i.e., "temporal adaptation") are likely to originate from temporal changes in cochlear nonlinearities.

Over the years, two main hypotheses have been developed regarding the cochlear mechanisms that underly temporal adaptation of the CEOAE level-curve. One school of thought proposes OHC adaptation to underly the effect (Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Verhulst et al., 2008), whereas another claims that an adaptive nonlinearity may not be needed. Instead, altered compression

properties would be created due to temporal overlap of BM impulse responses that would create an increased input to a static (i.e., time-invariant) nonlinearity in a given temporal window (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Hine and Thornton, 2002; Harte et al., 2005). To test the validity of the two proposed hypotheses, a new research approach is needed. Experimental CEOAE data are not able to directly link dynamic changes in *local* BM input/output functions to the CEOAE level-curve because the CEOAE constitutes a broadband cochlear response, stemming from various cochlear locations.

To further investigate the cochlear origins of temporal adaptation in CEOAEs, a nonlinear time-domain model of the cochlea is developed in **Chapter 4**. This model is able to simulate otoacoustic emissions and to reproduce features of published animal BM impulse response data. Details on the model implementation are explained in Chapter 4, and the performance of the model is evaluated considering key features of cochlear processing. The developed model allows for an investigation of the compression properties of simulated CEOAEs while, at the same time, changes in local BM I/O functions can be monitored.

Chapter 5 employs the developed time-domain model to investigate the cochlear origins of temporal adaptation in CEOAEs. To investigate the hypothesis that states that a time-dependent cochlear nonlinearity is not needed to account for temporal adaptation in the CEOAE level-curve, the model is implemented with a static (i.e., time-invariant) nonlinearity. Level-dependent features of CEOAEs are investigated in this chapter, together with the cochlear origin of temporal adaptation of the CEOAE level-curve.

Chapter 6 summarizes the main findings in this thesis. The main hypotheses behind temporal adaptation of the CEOAE level-curve are discussed in light of the experimental and simulated results presented. The implications of the results with respect to the existence of dynamic processes in cochlear nonlinearities are addressed.

2

Temporal Suppression and Augmentation of Click-Evoked Otoacoustic Emissions [1]

This study investigates temporal suppression of click-evoked otoacoustic emissions (CEOAEs), occurring when a suppressor click is presented close in time to a test click (e.g. 0–8 ms). Various temporal suppression methods for examining temporal changes in cochlear compression were evaluated and measured here for seven subjects, both for short- and long-latency CEOAEs. Long-latency CEOAEs (duration $>$ 20 ms) typically indicate the presence of synchronized spontaneous otoacoustic emissions (SSOAEs). Temporal suppression can only be linked to changes in CEOAE compression if the suppressor click affects CEOAE magnitude only. Phase changes induced by the suppressor click were shown to bias suppression in two ways: (i) when a specific asymmetric measurement method was used and (ii) when synchronization between the CEOAE and the click stimuli was incomplete. When such biases were eliminated, temporal suppression and augmentation (the opposite effect) were observed and shown to be subject dependent. This indicates that the nonlinearity underlying temporal suppression can work in a more (i.e., suppressed) or less (i.e., augmented) compressive state, depending on the inter-click interval and the subject under test. Temporal suppression was shown to be comparable for CEOAEs and SSOAEs, indicating similar underlying cochlear nonlinear mechanisms. This study contributes to a better understanding of the temporal properties of cochlear dynamics.

[1] This chapter was published as Verhulst et al. (2008)

2.1 Introduction

The active mechanisms in the cochlea, responsible for sharp tuning of the auditory filters and dynamic range compression, are also believed to give rise to otoacoustic emissions (OAEs), as first demonstrated by Kemp (1978). The majority of normal hearing adults (98 %) have measurable click-evoked otoacoustic emissions (CEOAEs) with emission spectra containing several discrete frequencies, known as dominant frequencies (Probst et al., 1991). These dominant frequencies can be categorized into two groups based on the generator mechanism behind them. The dominant frequencies in the first group are formed by oscillations at specific locations along the basilar membrane that synchronize to transient stimuli (Probst et al., 1991). The evoked oscillations are sustained for about 20 ms after the click onset and are referred to as *short-latency* CEOAE components (Probst et al., 1991). The second group of dominant frequencies in the CEOAE is linked to spontaneous otoacoustic emissions (SOAEs). SOAEs can be measured in the ear-canal even when no stimulus is presented and are often thought of as a consequence of particular locations along the cochlea that oscillate spontaneously because of a feedback process in the active cochlea (Eguiluz et al., 2000). Alternatively, SOAEs have been described as amplitude stabilized standing waves that are produced by the cochlea (Shera, 2003). SOAEs can synchronize to a presented click stimulus, so that they become synchronised SOAEs (SSOAEs; Wilson, 1980). These SSOAEs form the second group of dominant frequencies in the CEOAE and have durations longer than 20 ms. They are therefore referred to as *long-latency* CEOAE components (Probst et al., 1991). A CEOAE thus consists of a short-latency component that is broadband and a long-latency component that is related to SSOAEs.

CEOAE levels grow linearly with stimulus level for input levels below 20–30 dB and saturate with a rate of about 0.33 dB/dB for higher stimulus levels (Kemp, 1978; Probst et al., 1991). This compressive CEOAE level-curve is an important feature of the underlying CEOAE generator mechanisms and is very similar to the compression observed in the basilar membrane velocity for medium to intense sound inputs (Robles et al., 1986). Both basilar membrane and CEOAE compression are thought to have similar origins in the cochlear amplifier (Gold, 1948; Davis, 1983)

and have been investigated in the past by means of suppression in CEOAEs (Kemp and Chum, 1980; Tavartkiladze et al., 1994; Lina-Granade and Collet, 1995; Kevanishvili et al., 1996; Kapadia and Lutman, 2000b; Hine and Thornton, 2002). Suppression is defined as the *amplitude reduction* that can be observed in the CEOAE when a so-called suppressor click is presented close in time (e.g. 0–8 ms) before a test click (Kemp and Chum, 1980). Suppression is commonly considered to be related to the CEOAE level-curve, and could thus give a measure for investigating temporal aspects of cochlear compression by investigating the variation of suppression when shifting the suppressor click in time. Historical studies reported that suppression varied systematically with the timing and the level of the suppressor click, being greatest for suppressor clicks presented 2–4 ms before the test click (Kapadia and Lutman, 2000b; Hine and Thornton, 2002). Since suppression showed a time dependency related to the duration of the inter-click interval (ICI) between the suppressor and test click, the suppression effect was also referred to as temporal suppression. Temporal suppression has been modeled by means of a static nonlinearity (Kapadia and Lutman, 2000a; Harte et al., 2005). This static nonlinearity represented the CEOAE level-curve, linear for inputs below 30 dB and compressive with a constant factor between 0.33 and 0.5 dB/dB for inputs above this level. Temporal suppression was accounted for by varying the input to the static nonlinearity, depending on the ICI, and this is illustrated in Fig. 2.1.

In the reference condition (trace A), the input to the static nonlinearity consisted of an idealized cochlear output at a certain location along the basilar membrane. The cochlear output to a click stimulus was modeled as 10 cycles of a 1 kHz symmetric tone burst. When presenting two clicks temporally close to each other, the cochlear outputs to the individual clicks overlapped, leading to variations in the resulting signal presented to the static nonlinearity (trace B, C and D). The ICI-based changes in the simulated cochlear outputs led to different outputs of the static nonlinearity, reflected in different suppression levels. This model simulated suppression relatively well; maximal suppression was obtained for coincident clicks and any temporal separation, smaller than the duration of a single cochlear output, resulted in decreased suppression (Kapadia and Lutman, 2000a). The model could not account for maximal suppression at ICIs other than zero, and therefore Harte et al. (2005) introduced a delay into

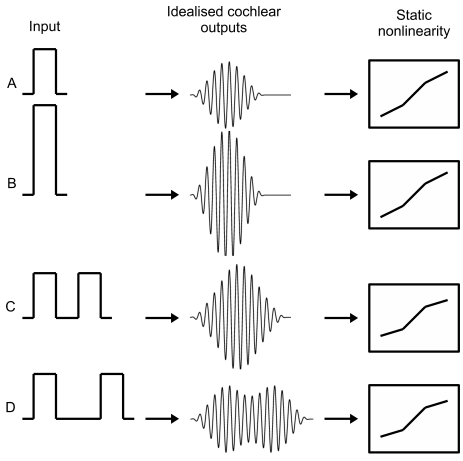


Figure 2.1: Illustration of the input signals that were presented to the static nonlinearity described in Kapadia and Lutman (2000a). Trace A shows the reference situation, where the idealized cochlear output to a click was given by 10 cycles of a 1 kHz symmetric tone burst. Traces B, C and D show cochlear outputs to double click configurations with varying ICI (resp. A: 0 ms, B: 3 ms and C: 6 ms). The cochlear outputs in B, C and D were obtained by summing time-shifted cochlear outputs to single clicks. The static nonlinearity was constant across the four traces, emphasizing that the only changes in the simulated suppression values of Kapadia and Lutman (2000a) were due to ICI-related changes of the input signals.

the model. This improved the model such that the experimental data of Kapadia and Lutman (2000b) was modeled more accurately. A discrepancy exists in the experimental studies about the observation of augmentation (Kemp and Chum, 1980; Tavartkiladze et al., 1994; Lina-Granade and Collet, 1995; Kevanishvili et al., 1996; Kapadia and Lutman, 2000b; Hine and Thornton, 2002), which is defined as the *increase* in CEOAE level when presenting a suppressor click close in time (e.g. 0–8 ms) before a test click. Augmentation cannot be accounted for when using a model based on the CEOAE level-curve, implemented as a static compressive nonlinearity. Changing the magnitude of the input to this static nonlinearity can change the amount of suppression that can be modeled, but can never create augmentation. This is the reason why the nonlinearity proposed by Kapadia and Lutman (2000a) can model temporal suppression relatively well, but cannot account for augmentation. In a companion study, Kapadia and Lutman (2000b) did not consider this as a problem, as their own as well as Kemp and Chum (1980)'s experimental data did not show

any augmentation. In contrast, the studies of Tavartkiladze et al. (1994) and Hine and Thornton (2002) clearly observed augmentation in their data. The latter studies did not present a modeling framework, but their data suggests that the static nonlinear models of Kapadia and Lutman (2000a) and Harte et al. (2005) would need to be extended to account for this augmentation.

Augmentation is essential for a full understanding of the characteristics of temporal suppression and thus requires further attention. The experimental conditions under which augmentation occur have not been clarified yet and were investigated in the present study. Different measurement paradigms were used in the past to assess temporal suppression and this might have introduced ambiguities when comparing temporal suppression across studies. In the present study, these earlier proposed measurement paradigms were applied on the same pool of subjects. The aim was to investigate why the results obtained by different paradigms differ and to test the validity of the paradigms in describing changes in cochlear compression. If augmentation is found to be due to methodological errors, then the proposed models of Kapadia and Lutman (2000b) and Harte et al. (2005) provide a good basis for describing temporal suppression. If augmentation is found to be linked to cochlear compression, then an extension to the existing models is required to account for suppression as well as augmentation.

During the present study, a distinction was made between temporal suppression of the short- and long-latency CEOAE. This allowed for an investigation of the generator mechanisms behind SSOAEs and CEOAEs. Other large studies have only investigated temporal suppression for evoked CEOAEs as they did not consider subjects with SOAEs (Kapadia and Lutman, 2000b; Hine and Thornton, 2002). It is hypothesized here that suppression would continue for the long-latency CEOAE if SOAEs and CEOAEs originate from identical cochlear nonlinearities inside the cochlea.

2.2 Materials and Methods

2.2.1 Recording CEOAEs

The stimuli were generated in Matlab and sent to a soundcard (RME FireFace 800 A/D-D/A converter) via the open source software pa-wavplay[1]. The signal levels were controlled via a DT-PA5 programmable attenuator and the stimuli were presented to the test subjects via an ER-2 probe. Recordings were made using an ER-10B low noise microphone, and bandpass filtered between 0.6 and 5 kHz (analog Rockland 852 HI/LO filter). The recorded analog signals were converted and stored digitally. Test subjects were screened to ensure that all had normal hearing (with a loss less than 15 dB HL across the frequencies considered in the audiogram). All the subjects were required to lie down in a soundproof booth with the probe in the left ear-canal.

2.2.2 Experimental paradigm

Three experiments were run in this study. In all of them, click stimuli were used to activate a broad range of frequency regions on the basilar membrane, leading to a spectrally rich CEOAE recording. Two clicks were presented, the first click was labeled as the suppressor click and the latter one as the test click. Both were presented at peak-equivalent sound pressure levels (peSPL) above 50 dB. This level was chosen to ensure that the response would be in the nonlinear compressive region of the level-curve for CEOAEs (Kemp, 1978). Seven subjects were tested and screened to have particularly strong CEOAEs with a long duration and with many CEOAE components. The recording window for the CEOAE was 49 ms long to allow an investigation of the long-latency click response associated with SOAEs. There was a possibility that the SSOAEs to the click stimuli of one recording window continued to ring into the next recording window. The scope of the temporal suppression effect was $\sim$10 ms whereas the recording window length was considerably longer than that. Therefore, it was expected that the window duration had no influence on the temporal suppression

[1] pa-wavplay contains a set of dll's allowing simultaneous play-rec for multiple channels. The code was written by Matt Frear (Univ. of Western Sydney) and uses the open source portaudio API. The open source code of pa-wavplay can be found on http://sourceforge.net/projects/pa-wavplay (last visited: 19^{th} of May 2008)

effect. Secondly, the existing SSOAEs at the start of a new window could have had an influence on the generation of the SSOAEs after the first click in that window. This means that the starting phase of the SSOAE after the click could have been altered by the relative phase of the SSOAE to the start of the click. To rule out this possibility, the influence of the recording window duration on to the recorded short- and long-latency CEOAE was investigated for one subject. Ten window durations between 28.3 and 50 ms were tested, and there was no apparent influence of the window duration onto the recorded CEOAE. This means that the relative phase of the SSOAE at the start of a click stimulus is of little or no importance for the generation of the SSOAE to this click stimulus. For every measurement point, 2000 click recordings were made to obtain an averaged click response with a reasonable signal-to-noise ratio.

In experiment 1, the suppressor and test clicks were separated in time, with inter-click intervals (ICIs) of 0.2, 0.33, 0.5, 1, 2, 3.33, 5, 6, 7 and 8 ms. These ICIs were chosen to reproduce experimental conditions from similar studies (Kapadia and Lutman, 2000b; Hine and Thornton, 2002). The ICIs of 7 and 8 ms were added to the set of ICIs used by Hine and Thornton (2002) to extend the scope of observation. The suppressor and test clicks had constant levels of 65 dB peSPL during experiment 1.

In experiment 2, the synchronization properties of the CEOAE components to the presented click stimuli were tested. Stimulus-synchronization, as proposed by Wilson (1980), is necessary for a correct interpretation of temporal suppression in terms of cochlear compression. When stimulus-synchronization is absent, the suppressor measure consists of a suppressor-induced *magnitude* and *phase* component. Since the focus of this study is to investigate temporal changes in cochlear compression and thus the magnitude changes of the CEOAE, an attempt was made to remove the suppressor-induced phase component from the suppression measure. Suppressor-induced phase changes could be due to an undefined temporal mechanism in the CEOAE generators, but this was not further investigated in this study. The phase component in the suppression measure can be omitted when assuring stimulus-synchronization of the CEOAE components to the presented click stimuli. The presence of stimulus-synchronization was tested for all seven subjects. Furthermore, two representative subjects from experiment 1 were tested more extensively to find good synchronization conditions. The ICI in the latter case was kept fixed at 2 ms, with the test-click level

(L_T) fixed at 65 dB peSPL and the suppressor-click level (L_S) varied between 50 and 75 dB peSPL in 5-dB steps.

In the third experiment, experiment 1 was repeated for the two exemplary subjects. All ICIs were remeasured for suppressor- and test-click levels of $L_S = 50$ and $L_T = 65$ dB peSPL, respectively, to make sure that all CEOAEs synchronized well with the presented click stimuli. The suppression measure in experiment 3 was not biased by a suppressor-induced phase component and could thus correctly investigate cochlear compression by means of suppression in CEOAEs.

Two methods were used to measure temporal suppression caused by a suppressor click presented prior to a test click. Method 1 was based on Tavartkiladze et al. (1994) and method 2 was introduced by Kemp and Chum (1980).

Method 1

This method was introduced by Tavartkiladze et al. (1994), and adopted by Kevanishvili et al. (1996) and Hine and Thornton (2002) in further studies. For each ICI, two responses were measured: (i) the *double-click response* consisting of a suppressor click (S-click) and a test click (T-click), and (ii) the *single-click* or *unsuppressed response* with the click positioned at the S-click of the double-click stimulus. The left panel in Fig. 2.2 shows a schematic of method 1 in which trace (a) and (b) represent the measured single- and double-click response, respectively. The single-click or unsuppressed response was subtracted from the double-click response (b)-(a) to yield the derived suppressed response (c). This response would equal a single-click response to the test click if CEOAEs had linear growth (i.e., if linear superposition held). Instead, a suppressed response is obtained that has an amplitude depending on the compression factor in the CEOAE level-curve and on the ICI separating suppressor and test click. Suppression in method 1 was defined as the difference in rms level between the derived suppressed response (c) and the unsuppressed response (d) (i.e., a CEOAE measured to a single click). This metric is positive when the derived suppressed response was smaller than the unsuppressed response. Negative suppression (i.e., when (c)>(d)) is referred to as augmentation.

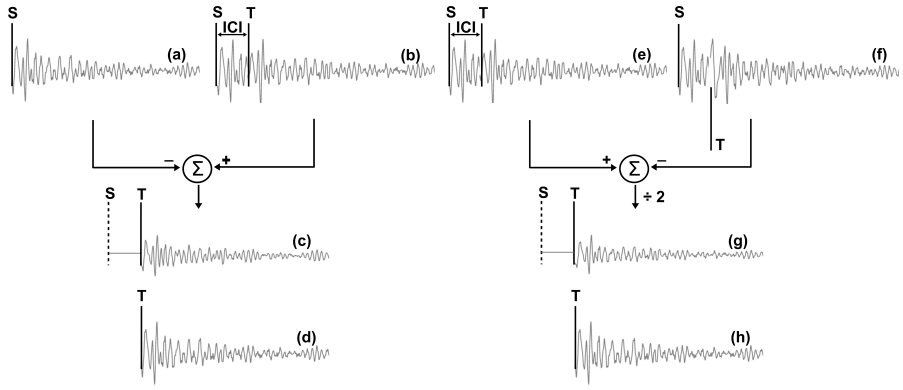


Figure 2.2: Schematic of method 1 (left panel) and method 2 (right panel). Left panel: trace (a) and (b) show a pair of CEOAE responses. (b) contains a S- and T-click separated by a certain ICI; (a) only contains a S-click. (c) shows the derived suppressed CEOAE obtained by aligning and subtracting (a) from (b). Trace (d) represents the reference unsuppressed CEOAE. Right panel: trace (e) and (f) show CEOAEs to double click stimuli; in (e) the S- and T-click have the same polarity; in (f) the T-click is inverted. The derived suppressed response in (g) is obtained by subtracting (f) from (e) and dividing the response by two. Suppression for method 1 and 2 is found as the difference between the rms level of the unsuppressed CEOAE (d), resp.(g) and the derived suppressed CEOAE (c), resp. (h) and is by definition positive when (c) $<$ (d), resp. (g) $<$ (h).

Method 2

Method 2 was introduced by Kemp and Chum (1980), and adapted by Lina-Granade and Collet (1995) and Kapadia and Lutman (2000b). There were some differences in method 2 across these studies. However, Kapadia and Lutman (2000a) showed that these did not affect the results. The specific version of method 2 presented here was the method introduced by Kapadia and Lutman (2000a). In this method, two double-click responses were measured for each ICI. The first click pair contained the suppressor and test click with the same polarity, and the second click pair contained the suppressor with positive and the test click with inverted polarity. The corresponding responses are indicated in the right panel of Fig. 2.2, as trace (e) and (f), respectively. The derived suppressed CEOAE in method 2 (g) is given by ((e)-(f))/2. When subtracting (f) from (e), the S-click has disappeared but the T-click has doubled, and this is compensated for by the division by 2. Suppression is defined as above, as the difference between

rms level of the derived suppressed response in (g) and the unsuppressed response in (h) (i.e., a CEOAE measured to a single click).

2.2.3 Post-processing

An artifact-rejection template was applied to remove ∼10% of the noisiest epochs, to increase the signal-to-noise ratio. The click responses were then averaged and the early-latency linear response, associated with the ear-canal and middle-ear transfer function, was removed. To determine this separation point from the early-latency response respective to the CEOAE, two single-click responses were measured at 65 and 75dB peSPL. Both responses were normalized and plotted on top of each other. The linear early-latency components sat atop, and the CEOAE components diverged due to the CEOAE compression. This technique is similar to the NL subtraction paradigm of Kemp et al. (1986). However, it has the advantage of just giving a time point to separate the CEOAE-dominated response from the early-latency response for a given subject, without actually affecting the CEOAE time series. The linear response in the click recordings disappeared between $\tau_s = 5$ and 6 ms after the last click onset for all subjects. Further analysis was carried out on the CEOAE component, set to start $\tau_s = 6$ ms after the last click onset for all subjects. Rms levels of the unsuppressed and suppressed CEOAE were calculated for 3-ms time frames between $\tau_s = 6$ and 36 ms after the last click onset via:

$$L_{rms}(t_p) = 20 \log_{10} \left[\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_{x_i})^2 \right]^{1/2} \tag{2.1}$$

with $N = 3$ ms $\cdot f_s$ and N being the number of samples in the 3-ms analysis window. x_i represents the i^{th} samples of the CEOAE and μ_{x_i} stands for the mean of the CEOAE within the considered 3-ms window. A value for L_{rms} is found for the time frame t_p of the CEOAE. Rms levels were furthermore calculated for the short- and long-latency CEOAE instead of for all 3-ms time frames, to facilitate a comparison across subjects. In this case, the rms values were calculated between $\tau_s = 6$ and 15 ms for the short-latency CEOAE, and between $\tau_s = 20$ and 36 ms for the long-latency

CEOAE. Levels of suppression were calculated as:

$$S(t_p) = L_{unsup}(t_p) - L_{sup}(t_p) \tag{2.2}$$

where $L_{unsup}(t_p)$ and $L_{sup}(t_p)$ represent the rms levels ($L_{rms}(t_p)$) of a certain time frame of respectively the unsuppressed and suppressed CEOAE. Negative suppression is also referred to as augmentation and occurs when $L_{unsup}(t_p) < L_{sup}(t_p)$.

2.3 Results

The results of the three experiments are presented in the following. For experiment 1, the average data are presented and discussed. This is followed by an evaluation of the measurement methods based on the data of a single representative subject. Stimulus-synchronization is found to be of critical importance for a correct interpretation of the suppression measure. Experiment 2 describes the situations under which stimulus-synchronization of the CEOAE to the click stimuli is present. Experiment 3 shows a repetition of experiment 1 where the stimulus levels where chosen such that stimulus-synchronization was present.

2.3.1 Experiment 1

Average data

Suppression levels for 3-ms time frames of the CEOAEs were averaged across seven subjects and are shown in Fig. 2.3 as a function of the inter-click interval (ICI) for both methods. The left panels of Fig. 2.3 represent suppression in the short-latency time frames (6–21 ms) of the CEOAE. The right panels show suppression in the long-latency (21–36 ms) time-frames of the CEOAE. The upper panels represent the suppression levels obtained with method 1 and the lower panels show suppression for method 2. Suppression (i.e., $S > 0$) was observed for the short- and long-latency CEOAE for ICIs shorter than 5 ms and for both methods presented in Fig. 2.3. Furthermore, method 1 showed augmentation (i.e., $S < 0$) for the short- and long-latency CEOAE for ICIs longer than 3.33 ms.

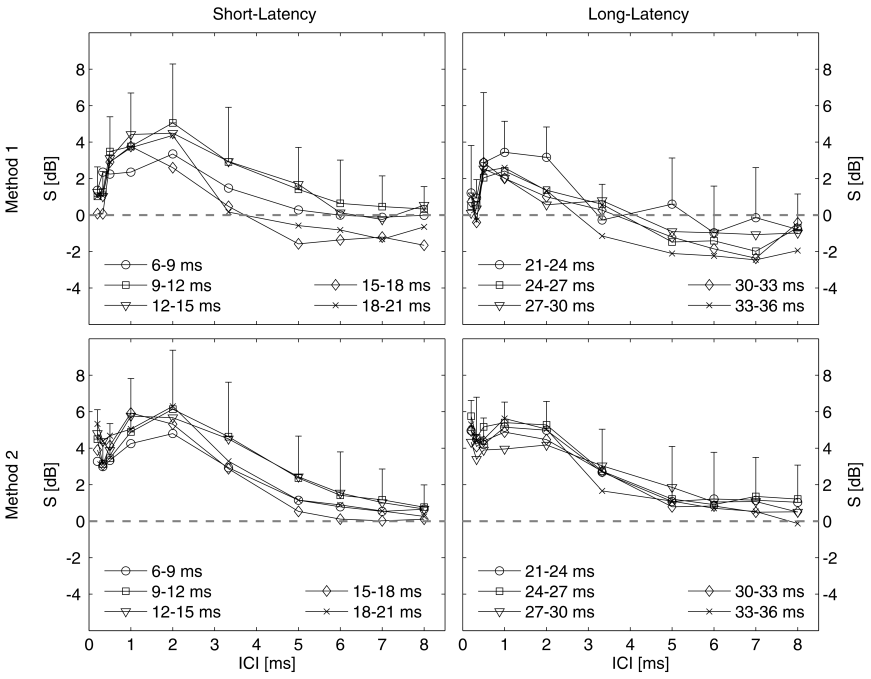


Figure 2.3: Rms levels of suppression (S) as a function of ICI, plotted for different short-latency time frames after the last click onset (left panels) and long-latency time frames (right panels), as an average over seven test subjects. Method 1 was used to obtain the suppression levels in the top panels and method 2 for the bottom panels. The standard deviation across subjects was plotted on the suppression levels of the 9–12 ms time frame (left panels) and of the of the 21–24 ms time frame (right panels) and representative for all time frames within the panel. The gray dashed line represents zero suppression and augmentation. L_S and L_T were 65 dB peSPL.

The short-latency suppression (left panels in Fig. 2.3) increased with ICI duration for all time frames until maximal suppression was reached at an ICI of 1–2 ms, after which suppression decayed with increasing ICI to a constant level between -2 and 1 dB. Maximal suppression was found at ICIs of 1–2 ms for all time frames, with values of 2–5 dB for method 1, and of 4–7 dB for method 2. The short-latency suppression obtained by method 1 (upper-left panel in Fig. 2.3) is quantitatively comparable to the suppression reported by Hine and Thornton (2002), who found maximal suppression

of 4–5 dB at an ICI of 1 ms for stimulus levels of 65 dB peSPL and time frames between 7 and 17 ms. Qualitatively, the data of Tavartkiladze et al. (1994) and Kevanishvili et al. (1996) showed similar trends of suppression as a function of ICI. Augmentation of the CEOAE (i.e., $S < 0$) was observed using method 1 for ICIs between 5 and 8 ms and for time frames between 15 and 21 ms (upper-left panel in Fig. 2.3). This is in agreement with Tavartkiladze et al. (1994), who reported augmentation of the CEOAE for an ICI of 7 ms. However, Kevanishvili et al. (1996) and Hine and Thornton (2002) have not observed augmentation for the short-latency CEOAE in the double-click paradigm.

The short-latency suppression obtained with method 2 (lower-left panel in Fig. 2.3) is quantitatively in agreement with the data of Kapadia and Lutman (2000b). They reported maximum suppression at ICIs of 1–2 ms with values of 5–6 dB for click levels of 60 dB peSPL. Furthermore, the short-latency suppression obtained with method 2 is quantitatively comparable to the observed suppression in Kemp and Chum (1980) and Lina-Granade and Collet (1995). The general observation of suppression across ICI for the short-latency CEOAE was similar for both methods, although there were two clear differences. Firstly, the overall suppression curves of all time frames seemed to be offset by about 2 dB in the direction of more suppression for method 2. As a consequence, there was no augmentation of the short-latency CEOAE in method 2. Secondly, the difference between the suppression curves for different time frames was smaller for method 2 than for method 1.

Suppression levels of the long-latency CEOAE are shown in the right panels of Fig. 2.3 for both methods. Temporal suppression was generally found for the long-latency CEOAEs, indicating that temporal suppression affected the SSOAEs. For method 2, suppression across ICI showed very similar trends for the long-latency CEOAE as for the short-latency CEOAE. For method 1, two differences were observed between short- and long-latency suppression. Firstly, maximum suppression occurred at an ICI of 0.5 ms for the long latencies, representing a small shift towards smaller ICIs for maximal suppression. Secondly, augmentation was observed for all ICIs above 3.33 ms in the long-latency CEOAE. Former studies have not investigated suppression of long-latency CEOAEs.

When comparing suppression obtained by the different measurement methods,

two differences were noticed. Method 1 showed a greater difference between the amounts of suppression across ICI than method 2. This could indicate that, using method 1, suppression affected certain parts of the CEOAE waveform differently than others or that this method was more affected by noise than method 2. Method 1 showed decreasing suppression levels with latency for ICIs greater than 3.33 ms, manifested as a general offset of about 2 dB towards lower values for the long-latency CEOAE than for the short-latency CEOAE. Method 2, on the other hand, did not show apparent differences of suppression with latency. The latter observation would indicate that suppression affected SOAEs and CEOAEs similarly, whereas the results obtained with method 1 would argue against this. The observed differences in suppression related to the methods can thus lead to different interpretations of the temporal suppression effect and this required a closer examination of the definitions of suppression in both methods. This evaluation was conducted in section 2.3.1.

The standard deviation across subjects in the data in Fig. 2.3 was quite large. Table 2.1 reports suppression obtained with method 2 for the short- and long-latency CEOAEs of all seven subjects that participated in the study. It can be observed that, for a single ICI, the levels of suppression varied by about 4 dB (i.e., two standard deviations) across subjects. The results in Table 2.1 also indicate that the dependency of suppression on ICI was similar across subjects. However, the exact levels of suppression were subject dependent. Thus, some subjects show only suppression and others show both suppression and augmentation (Table 2.1). In order to clarify the conditions under which augmentation occurs, a further analysis based on the data of single subjects, instead of on the average suppression data, was undertaken.

The goal was to investigate whether augmentation could be a consequence of the methodology used here or whether it was due to a lesser amount of compression in the auditory system. A representative subject of the pool of seven subjects was chosen to conduct this analysis for. This particular subject showed suppression for ICIs below 3.33 ms and augmentation for the longer ICIs, and appears thus well suited to investigate augmentation.

Table 2.1: Suppression levels (S) in dB for the short- and long-latency CEOAE for 10 ICIs and all seven subjects. S was calculated with method 2 between $\tau_s = 6$ and 15 ms for the short-latency CEOAE (SL) and between $\tau_s = 20$ and 36 ms for the long-latency CEOAE (LL). The bold figures represent suppression values for the situation where the recorded double-click responses with method 2 were directly out of phase. This indicated that stimulus-synchronization was present for the CEOAEs, used to calculate S. 'S1' represents subject 1 taken from Fig. 2.4. The bottom two lines show the standard deviation across subjects for the short- and long-latency CEOAE for a single ICI.

ICI [ms]	0.2	0.3	0.5	1.0	2.0	3.3	5.0	6.0	7.0	8.0
'S1' SL	2.6	2.9	2.5	3.8	3.8	1.6	**0**	**0**	**-0.6**	**-0.6**
LL	1.6	1.4	1.9	3.8	2.1	**-1.2**	**-3.1**	-2.3	-1.7	**-2.6**
'S2' SL	5.4	4.5	4.4	5.6	6.6	4.5	**1.9**	**1.2**	**0.6**	**-0.2**
LL	9.6	7.1	6.4	7.9	6.2	3.4	**1.3**	**1.6**	**1.6**	**1.4**
'S3' SL	4.0	2.5	4.2	4.5	5.8	2.0	**-0.1**	**-0.5**	**0**	**0**
LL	2.8	0.5	3.8	3.4	5.0	0.7	**-2.1**	**-2.7**	-2.0	-1.6
'S4' SL	3.9	2.6	3.9	6.2	8.4	6.8	3.5	**2.3**	**2.7**	**2.6**
LL	2.7	4.7	3.7	3.0	5.0	4.1	5.0	**6.0**	**5.6**	**4.0**
'S5' SL	3.1	1.8	**2.5**	3.0	3.4	3.3	2.8	**2.3**	**1.7**	**1.2**
LL	6.0	4.3	**3.0**	4.3	5.6	5.1	3.1	2.4	**1.7**	**1.1**
'S6' SL	4.7	3.9	4.0	4.6	3.6	2.2	1.4	1.1	**1.0**	**1.1**
LL	4.2	2.9	4.4	3	1.3	1.6	1.2	0.5	**-0.4**	**-0.8**
'S7' SL	**4.4**	5.4	**3.3**	5.5	7.0	6.3	3.3	**1.6**	**0.6**	**0.2**
LL	7.7	7.2	**7.3**	8.5	7.0	4.1	3.5	1.5	**2.0**	**1.9**
Std dev SL	1.6	1.1	1.9	2.9	3.2	3.0	2.3	2.4	1.7	1.2
LL	1.7	2.3	1.5	1.4	1.6	2.4	3.0	2.6	2.4	2.0

Evaluation of the measurement methods

Fig. 2.4 shows suppression levels for 3-ms time frames of the CEOAE of one representative subject taken from the dataset in Fig. 2.3. It is observed from Fig. 2.4 that all four panels, representing both methods and both latency regions, show augmentation. This is in contrast to the presented average data in Fig. 2.3, where augmentation was only observed in the case of method 1. These results indicate that augmentation can indeed occur for both measurement methods and is not inherent to the usage of method 1. For method 1, the long-latency CEOAE showed a general offset to more augmentation across ICI compared to the short-latency CEOAE. For method 2, suppression of the short- and long-latency CEOAE had similar values up to an ICI of 3.33 ms. For longer ICIs, there was a tendency towards more augmentation for the long-latency CEOAE than for the short-latency CEOAE.

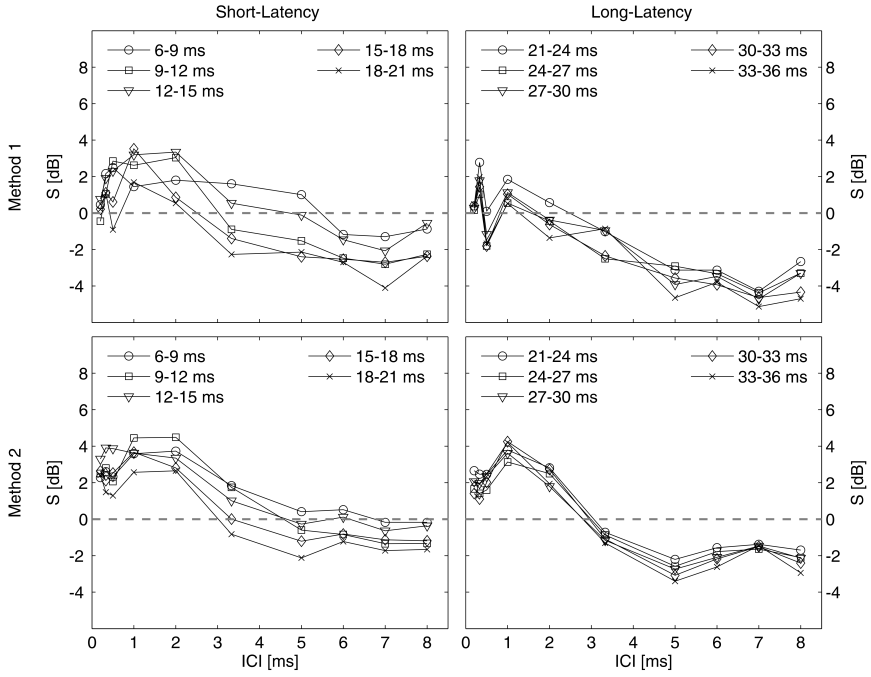


Figure 2.4: Rms levels of suppression (S) as a function of ICI, plotted for different short-latency time frames after the last click onset (left panels) and long-latency time frames (right panels), for one representative subject. Method 1 was used to obtain the suppression levels in the top panels and method 2 for the bottom panels. The gray dashed line represents zero suppression and augmentation. L_S and L_T were 65 dB peSPL.

The analysis of suppression for this single subject revealed differences that were related to the method, especially when comparing the long-latency suppression for both methods. Fig. 2.5 shows a selection of the derived suppressed CEOAEs, required to obtain suppression for subject 1 using method 1 (gray traces) or using method 2 (black traces). Suppression was largest for an ICI of 2 ms because the derived suppressed response for this ICI (trace 3) was smallest compared to the unsuppressed CEOAE (top trace). Augmentation of the derived suppressed waveforms was observed when comparing the CEOAE amplitudes at ICIs of 5 and 6 ms (2 bottom traces)

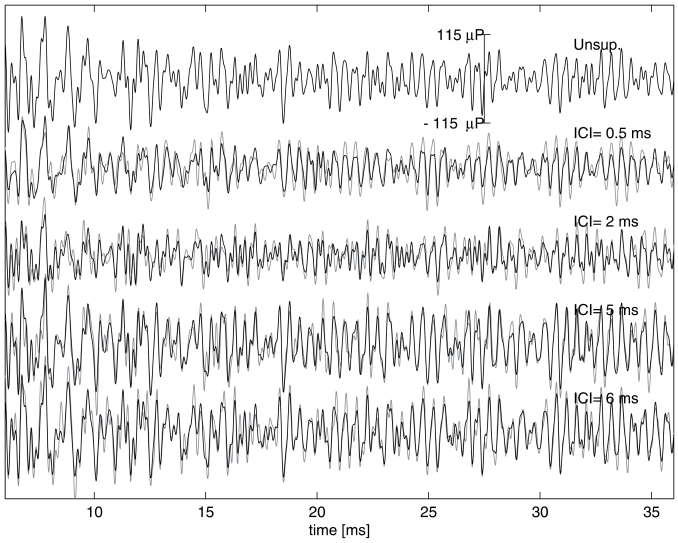


Figure 2.5: The derived suppressed responses of subject 1 for ICIs 0.5, 2, 5 and 6 ms. The gray traces show the derived suppressed responses obtained by method 1; the black traces were derived with method 2 and plotted on top of the gray traces. The top trace represents the unsuppressed response. Subtraction of the level of derived suppressed response from the level of the unsuppressed responses led to the suppression levels in Fig. 2.4.

to the unsuppressed CEOAE. When looking at the derived suppressed responses for both methods at a certain ICI, there were differences between the derived suppressed CEOAEs. Overall, the derived suppressed CEOAEs obtained by method 1 (gray traces) had larger amplitudes than the derived suppressed responses obtained by method 2 (black traces). This observation corresponds to the suppression presented in Fig. 2.4 that showed a tendency towards more augmentation for method 1 compared to method 2, especially for the long-latency CEOAE.

The deviating suppression results in Fig. 2.4 were due to differences in the derived suppressed responses for both methods and the cause of these differences should thus be found in the methods that were adopted to derive them. To investigate this, first, a simulation was run to investigate whether artifacts in the measurement methods could have been the basis of the differing suppression results. Secondly, the synchronization

of the CEOAE components with the click stimuli was tested. If there is a lack of synchronization, then phase differences could bias the suppression measure, leading to an incorrect interpretation of suppression in terms of compression. Wilson (1980) showed that stimulus-synchronization of the OAE occurs for click stimuli, but it has not been investigated whether this synchronization can also be found when two clicks are presented close in time to one another.

The simulation that investigated the derivation of the suppressed CEOAEs focused on the SSOAEs that appear as tonal components in the measured CEOAE. In the past, SOAEs have been modeled by limit-cycle oscillators (Eguiluz et al., 2000), which verifies that pure tones are suited as a simplified representation of SOAEs. The SOAEs were modeled to have a fixed magnitude response, so that the only suppression observed was due to artifacts in the measurement methods. The SOAE components were assumed to be perfectly synchronized to click stimuli, as commonly assumed in the literature (Probst et al., 1991). Stimulus-synchronization means that the modeled SOAEs would interact with click stimuli by resetting their phase upon stimulation, giving rise to synchronized SOAEs (SSOAEs) (Wilson, 1980).

Fig. 2.6 shows a representation of the derivation of the suppressed response for a SSOAE modeled as a pure tone. The derived suppressed responses of the modeled

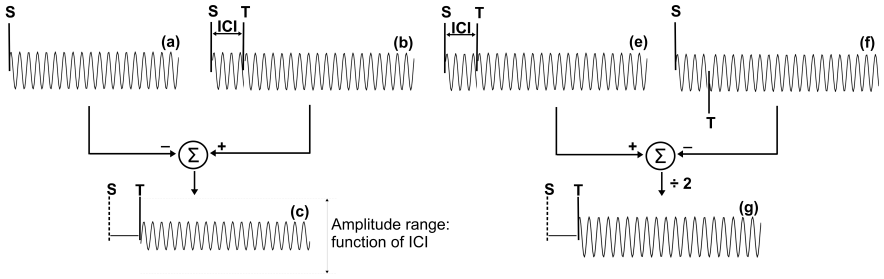


Figure 2.6: Representation of the methodology for a certain SSOAE component: method 1 (left panel) and method 2 (right panel). The stimulus in (b) contains a S- and T-click separated by a certain ICI; the stimulus in (a) only contains a S-click. The derived suppressed response in (c) is obtained by aligning the S-click in (a) with the S-click in (b) and subtracting (a) from (b). The amplitude of (c) varies with ICI due to phase differences. Right panel: trace (e) and (f) show SSOAEs to double click stimuli; in (e) the S- and T-click have the same polarity; in (f) the T-click is inverted. The derived-suppressed response in (g) is obtained by subtracting (f) from (e) and dividing the response by two. The derived suppressed response (g) is identical to the unsuppressed response, when there is no compression in the system.

SOAE were plotted in trace (c) and (g) in Fig. 2.6, respectively, for method 1 and 2. The amplitude of the derived suppressed response in method 1 varied according to the ICI. The alignment and subtraction of the SSOAE to a single-click stimulus (a) from the SSOAE to a double-click stimulus (b) introduced a phase component related to the ICI, under the assumption that the phase of the SSOAE reset upon click stimulation. The amplitude of the derived suppressed response equaled the unsuppressed response only when the ICI was a multiple of the natural period of the SSOAE component. In all other cases, the stimulus-synchronization of the SSOAE upon the second click led to a variation of the amplitude of the derived suppressed response. This amplitude variation would be interpreted as either suppression or augmentation, and is purely due to the relative phase. Method 2, in the right panel of Fig. 2.6, showed that the amplitude of the derived suppressed SSOAE always equaled the amplitude of the unsuppressed response, when no compression was introduced in the system. This was true for all possible ICIs, assuming that the SSOAE to the second click reset in phase correctly for a positive or negative polarity click (i.e., in phase for a positive polarity click and in anti-phase for a negative polarity click). Thus, only method 2 is theoretically able to associate suppression (as a magnitude effect) with compression in the CEOAE generator mechanisms, when there is compression in the physical system.

The differences in suppression between the methods for the long-latency CEOAE in Fig. 2.4 are likely due to the suppressor-induced phase bias on the levels for method 1. In the case of subject 1, this bias was towards more augmentation. The bias is subject dependent and could be either positive or negative, depending on the specific SOAE components present.

The phase-induced amplitude change of the derived suppressed SSOAE components due to method 1 could also have had an influence on the short-latency CEOAE, assuming that all dominant CEOAE-components synchronized with the click stimuli. The effect would have been smaller because the short-latency CEOAE was spectrally rich. Adding up the phase-induced amplitude fluctuations of a large amount of components interacting with the ICI would reduce the effect considerably. This was in agreement with the observation of a smaller deviation between suppression for the short-latency CEOAE obtained with method 1 compared to method 2 (right panels in Fig. 2.4).

Due to its independence of the ICI, method 2 offers the better choice for investigating CEOAE suppression and its link to the underlying CEOAE generator mechanisms. However, with method 2, suppression can only be related to compression in the CEOAE generator mechanisms when all CEOAE components are synchronized to both the suppressor and test click. This was examined in detail for all seven subjects in section 2.3.2. Furthermore, experiment 2 was conducted on two representative subjects to find optimal synchronization situations.

2.3.2 Experiment 2

Fig. 2.7 shows the CEOAEs of subject 1 that were used to obtain the derived suppressed CEOAEs for method 2 in Fig. 2.5. The CEOAEs to the double-click stimuli (corresponding to trace (e) in Fig. 2.2) are shown as the gray traces and the CEOAEs to the inverted polarity double-click stimuli (corresponding to trace (f) in Fig. 2.2) are shown as the black traces. In theory, perfect stimulus-synchronization would imply the CEOAE to the inverted click to be in anti-phase to the CEOAE to the regular click, i.e., the gray and the black traces would be in exact anti-phase, corresponding to a correlation coefficient of -1. There is an inherent noise floor present when measuring CEOAEs and this causes the correlation coefficient to have an absolute value of less than 1 for anti-phase signals. A visual inspection of Fig. 2.7 indicates that the gray and black curves appear to be in anti-phase for ICIs longer than 3.33 ms in case of the short-latency CEOAE ($<$20 ms). For the long-latency CEOAE ($>$20 ms), the gray and black curves tended to be in anti-phase for the ICIs of 3.33, 5 and 8 ms. The corresponding correlation coefficients had values between -0.8 and -0.5, indicating a tendency towards anti-phase behavior of these traces. Only for these ICIs, the suppression in Fig. 2.4 can be interpreted in terms of changes in compression. For all the other ICIs, the gray and black traces had some weaker correlation, and this means that the reported suppression for these ICIs cannot be linked to compression as they are partly biased by a phase component. The phase-induced suppression components were the cause for the observed augmentation of the long-latency CEOAEs for ICIs longer than 3.33 ms. Although there was one exception for an ICI of 8 ms. The double-click CEOAEs that led to this derived suppressed CEOAE were visually nearly perfect in anti-phase for this ICI (Fig. 2.7). This means that the augmentation observed

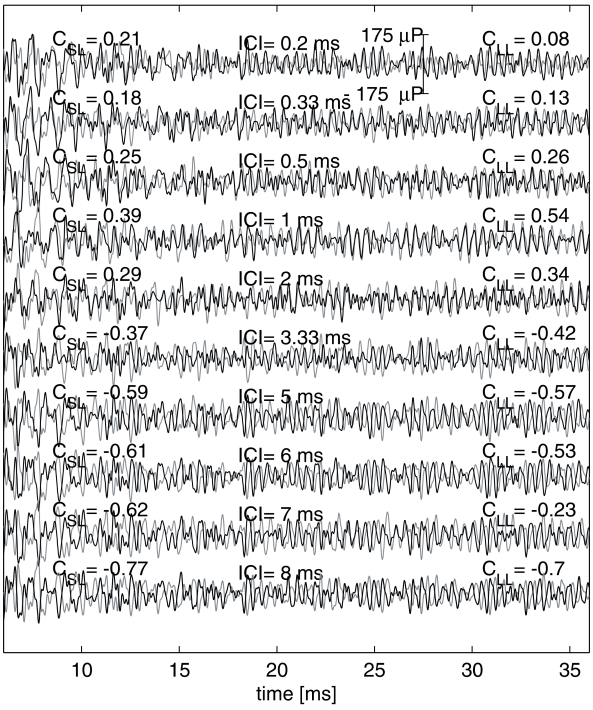


Figure 2.7: Stimulus-synchronization of the CEOAEs of subject 1 to the double-click stimuli in method 2. The gray traces show the CEOAEs after the double-click stimuli and the black traces show the CEOAE after the double-click stimuli with inverted polarity for the second click. The correlation coefficients between the gray and black trace are plotted on top of the short- and long-latency CEOAEs (C_{SL} and C_{LL} respectively). Stimulus-synchronization occurs when the gray and black traces are out of phase for a certain ICI (i.e., a correlation coefficient of -1). L_T and L_S were 65 dB peSPL.

in Fig. 2.4 for this single subject and for the ICI of 8 ms, was not biased by a phase component and can thus be related to changes in compression.

The suppression values for the other subjects that were calculated by method 2 in Table 2.1 were analyzed in terms of synchronization, and the values that were obtained under good stimulus-synchronization conditions were indicated in bold. Subject 2, 4 and 7 showed non phase-biased suppression between 1.6 and 5.6 dB at an ICI of 6

ms whereas subjects 1 and 3 showed non phase-biased augmentation values of -2.6 at an ICI of 8 ms and -2.7 dB at an ICI of 6 ms. This indicated that both suppression and augmentation were observed at the longer ICIs and that the underlying CEOAE generator mechanisms can thus both work in a more (i.e., suppressed) or less (i.e., augmented) compressive state depending on the subject under test and the ICI. There was a rather large spread of suppression across subjects (e.g. 4 dB for the long-latency suppression levels at an ICI of 8 ms) and this indicated that suppression and the underlying mechanisms were subject dependent. The results in Table 2.1 did not show sufficient data points where the stimulus-synchronization criterion was met to conclude that suppression is a robust measure to investigate compression in the cochlea. Only when the synchronization can be ameliorated, then suppression measured with method 2 could become a robust measure and this was tested in experiment 2.

Two subjects were tested to find level conditions where synchronization between the double-click stimuli and the CEOAEs occurred for all ICIs. Different levels of the suppressor click were presented while keeping the test click level constant. This was done for an ICI of 2 ms, chosen because suppression was shown to be large and the stimulus-synchronization to be poor for this experimental condition. The double-click CEOAEs (the test clicks had inverted polarities) of subject 1 are shown in Fig. 2.8 for a test level of 65 dB peSPL and varying suppressor levels. The phase behavior of the double-click CEOAEs in Fig. 2.7 changed drastically with suppressor level. When L_S equaled or exceeded L_T ($L_S \geq 65$ dB peSPL), no stimulus-synchronization was observed in the CEOAEs. The correlation coefficients indicated in phase behavior for these traces, which is opposite to what is required for good stimulus-synchronization. This finding implied that the CEOAEs did not always reset their phase upon stimulation by the second click, even though they synchronized to the first click. Method 2 can thus, for these level configurations, not be adopted to relate suppression or augmentation to compression. On the other hand, when L_S was smaller than L_T ($L_S \leq 65$ dB peSPL), the CEOAEs did synchronize to both clicks in the double-click stimuli, leading to improved anti-phase behavior between the traces. When L_S was 50 dB, the traces showed a near perfect stimulus-synchronization with a correlation factor that had an absolute value larger than 0.8.

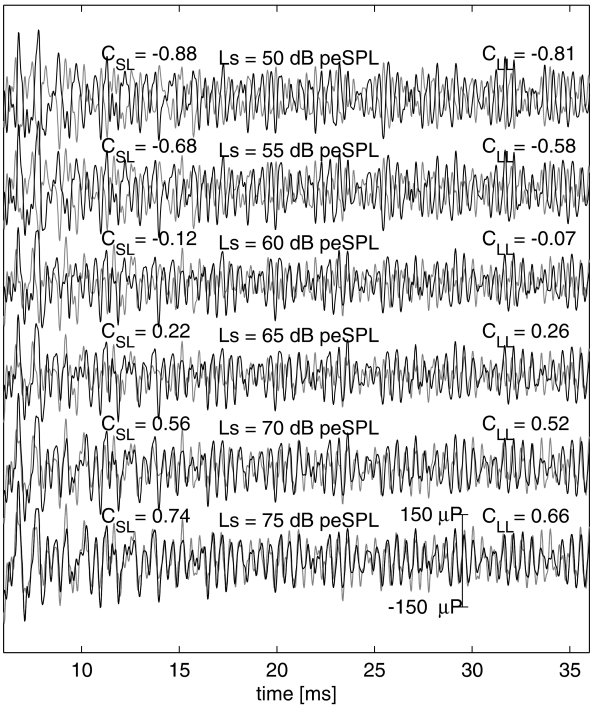


Figure 2.8: Stimulus-synchronization of the CEOAEs to the double-click stimuli of subject 1 for an ICI of 2 ms and method 2. The gray traces show the CEOAEs after the double-click stimuli and the black traces show the CEOAE after the double-click stimuli with inverted polarity for the second click. The correlation coefficients between the gray and black trace are plotted on top of the short- and long-latency CEOAEs (C_{SL} and C_{LL} respectively). L_T was 65 dB peSPL and L_S varied between 50 and 75 dB peSPL.

2.3.3 Experiment 3

Experiment 1 was repeated for subject 1 under the best possible stimulus-synchronization condition. L_S was set to 50 dB peSPL and L_T was kept at 65 dB peSPL, while all the ICI conditions were measured again. The recorded traces showed near perfect stimulus-synchronization for both the short- and long-latencies and the calculated suppression levels were plotted in Fig. 2.9. It was apparent from the suppression levels in Fig. 2.9 that the temporal suppression effect had a small dynamic range of

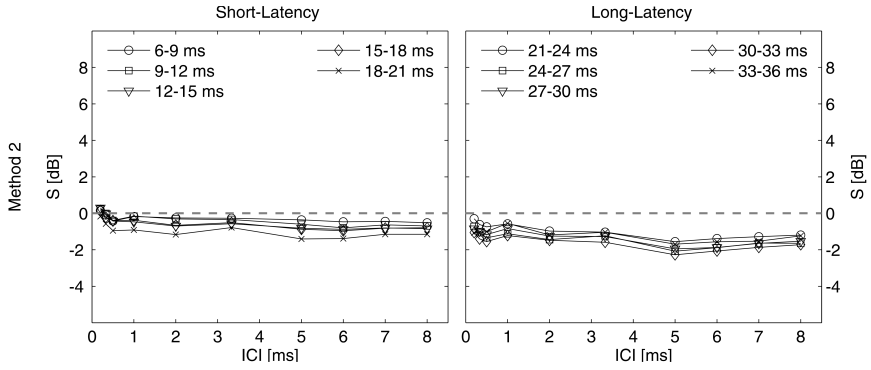


Figure 2.9: Rms levels of suppression (S) as a function of ICI, plotted for different short-latency time frames after the last click onset (left panels) and long-latency time frames (right panels), for one representative subject. Method 2 was used to obtain the suppression levels and the gray dashed line represents zero suppression and augmentation. L_S was 50 dB peSPL and L_T was 65 dB peSPL to ensure stimulus-synchronization.

about 2 dB. This was in agreement with the short-latency data of Kapadia and Lutman (2000b), that showed a dynamic range of 2 dB when the suppressor level was 10 dB lower than the test level. It was clear from Fig. 2.9 that there was no observation of suppression for subject 1, only augmentation. This is in contrast to the findings of Kapadia and Lutman (2000b), who did not observe augmentation when the suppressor click led the test click. The results in Fig. 2.9 showed a small tendency towards augmentation for the short-latency CEOAE and a larger offset to augmentation for the long-latency CEOAE. Augmentation increased in magnitude when increasing the time frame in the CEOAE.

The dynamic range of the whole temporal suppression effect decreased drastically from about 8 dB to 2 dB, when decreasing the suppressor level and thus assuring stimulus-synchronization of the CEOAE. An inspection of the long-latency suppression data of subject 1 in Table 2.1 showed that at the ICI of 8 ms, stimulus-synchronization was present during experiment 1. The 2.6 dB of augmentation reported for when $L_S = L_T = 65$ dB peSPL in Fig. 2.4 was larger than the 1.8 dB reported in Fig. 2.9 for the long-latency CEOAE at an ICI of 8 ms. The levels of augmentation decreased with decreasing suppressor-click level, which indi-

cated the existence of a trade-off between an effective suppressor and good stimulus-synchronization conditions. For equi-level stimuli, the dynamic range of temporal suppression was found to be 8 dB in Fig. 2.4. Since the suppression for this level configuration consisted of a magnitude- and phase-induced suppression component, the dynamic range of suppression related to cochlear compression was smaller than the observed 8 dB. It is estimated that, when compensating for phase-induced suppression, the dynamic range of the suppression effect related to cochlear compression would be reduced from 8 to 5 dB. This 3-dB reduction represents the worst-case bias that can be introduced by the lack of stimulus-synchronization (i.e., when traces (e) and (f) in Fig. 2.2 are in phase in stead of in anti-phase). Even if the dynamic range of temporal suppression was smaller than 5 dB when $L_S < L_T$, augmentation was observed in Fig. 2.9. This indicated that augmentation related to a lesser amount of compression in the CEOAE generator mechanisms caused by temporally spaced clicks presented to the ear is possible.

2.4 Discussion

2.4.1 Methodology for temporal suppression

Two classes of methods have been used in the past to quantify temporal suppression: the first method was based on a technique introduced by Tavartkiladze et al. (1994) and adopted by Kevanishvili et al. (1996) and Hine and Thornton (2002). In this method, a CEOAE to a single click was subtracted from a CEOAE to a double-click stimulus to get to the derived suppressed response. The second method differed from this by recording two CEOAEs to double-click stimuli with inverted clicks to derive the suppressed response and was introduced by Kemp and Chum (1980) and adapted by Lina-Granade and Collet (1995) and Kapadia and Lutman (2000b). This paper has demonstrated that suppression obtained by these methods consists of two main components: 1) suppression due to a changing amount of compression in the CEOAE generators and 2) a suppressor-induced phase component. This phase component was either related to the adoption of method 1 or to the absence of stimulus-synchronization of the CEOAE in method 2. The phase-induced suppressor component interferes with the

interpretation of temporal suppression in terms of compression. This phase-dependent suppression has been overlooked in the past, leading to incorrect suppression levels. The observed average suppression levels in Fig. 2.3 showed that our results, measured with both methods, were in agreement with the data obtained in previous studies. The differences between the used methods became apparent and further investigation showed that suppression obtained by method 1 was always obscured by a suppressor-induced phase component. This explained the bias towards augmentation in the experimental data obtained with method 1 compared to method 2. This bias depended strongly on the interaction between the natural period of certain CEOAE components and the ICI. It was therefore impossible to separate the phase-induced suppression from the suppression due to compression. The average suppression across seven subjects obtained with method 1 showed augmentation. However, this should be random and not representative for a pool of subjects as the bias introduced in method 1 is subject dependent. Individual subjects with varied dominant CEOAE components might yield different amounts of suppression, depending on the distinct phase interactions with the ICI. The suppressor-induced phase changes introduced by method 1 are more than likely the reason for the observation of augmentation at ICIs of 7–8 ms of the subjects used in the study of Tavartkiladze et al. (1994). Kevanishvili et al. (1996) did not show augmentation, once more confirming the subject dependency of temporal suppression when adopting method 1. When averaging suppression across subjects, it could be possible to rule out the suppressor-induced phase component because all the considered CEOAE components of all subjects would interact differently with the ICI. The phase component would then reduce to a noise on the underlying compression-related suppression. The number of subjects needed to achieve this phase compensation would need to be larger than the seven used here or in the study of Hine and Thornton (2002). It is concluded here that method 1 cannot be used to investigate compression in the CEOAE generator mechanisms.

Method 2 was robust to phase-induced suppression and therefore preferred for investigating compression in the CEOAE generator mechanisms. The average suppression obtained with method 2 in Fig. 2.3 did not show augmentation, in agreement with other studies that adopted method 2 (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b). Although method 2 was shown to be

phase blind, this was only true assuming that all CEOAE components synchronized to the double-click stimuli. Synchronization of the CEOAE components to the click stimuli was shown to be critical for a correct interpretation of the suppression levels obtained by method 2. It was shown in experiment 2 that stimulus-synchronization does not always occur. Table 2.1 and Fig. 2.7 showed that stimulus-synchronization was absent for most ICIs below 5 ms. A fixed phase delay was introduced in the CEOAE when presenting a second click shortly after a first click. This indicated that the CEOAE generation to the first click was quite robust for it to be disturbed by a second click stimulation during the first few milliseconds. During this time, a temporal effect took place that interrupted the CEOAE from resetting its phase and this effect was subject dependent. The exact relation between the phase delay, introduced by presenting a click shortly after another click, and the ICI has not been established. Stimulus-synchronization of the CEOAE is essential when investigating compression in the CEOAE generator mechanisms, but the lack of it does not necessarily mean that these phase interactions are artifactual. In fact, suppressor-induced phase changes can represent some undefined nonlinear temporal process, implying that one can no longer simply relate suppression with static compression. More complicated models (i.e., functional approaches) would need to be employed to begin to interpret such suppressor-induced phase changes. Some work has been done in investigating temporal nonlinearities in the CEOAE generators by the use of Voltera-Slice OAEs (Thornton et al., 2006; Lineton et al., 2008). However, due to the complexity of such analysis, little improved understanding of the underlying CEOAE generator mechanisms has been achieved to date.

Stimulus-synchronization properties of the CEOAE components were overlooked by the previous studies that adopted method 2 to investigate temporal changes in compression in the CEOAE generator mechanisms (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b). The importance of the synchronization of the OAE to external stimuli in investigating compression was shown in a related study by Neumann et al. (1997). They investigated suppression of OAEs, evoked by narrowband tone pulses, by suppressor tones and concluded that the *magnitude* of the OAE could not be altered by the suppressor tone. The observed reductions in OAE amplitude were solely due to phase interactions caused by synchronization

of the OAE to either the suppressor or the evoking stimulus. Thus, the suppression effect they discussed was purely artifactual in relation to compression. The absence of stimulus-synchronization can lead to an incorrect interpretation of temporal suppression as compression in the CEOAE generators and this is due to the phase-induced suppression components that are created when the synchronization is absent. This phase-induced suppression component must be taken into account when interpreting the suppression results obtained here and in previous studies (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b). The suppression data obtained in the latter studies still represent temporal suppression, only this effect is in most cases due to a combination between phase-induced suppression caused by poor stimulus-synchronization and compression-induced suppression in the CEOAE generator mechanisms. In this study, suppression was shown not to be biased by phase-induced suppression when stimulus-synchronization was present, so that in those cases suppression could be interpreted in terms of cochlear compression. Individual suppression data at the ICIs between 6 and 8 ms (Table 2.1) showed values between -2.6 and 6 dB that were related to compression in the CEOAE generator mechanisms. The observation of both augmentation and suppression indicated that the CEOAE generator mechanisms could both work in a less (augmentation) or more (suppression) compressive manner, depending on the subject and ICI.

In experiment 3, the suppressor click had lower levels than the test click, making sure that the CEOAE always synchronized to the click stimuli for all ICIs. The suppression obtained under these conditions for one subject (Fig. 2.9) showed that the dynamic range of temporal suppression had reduced drastically to about 2 dB. This was not unexpected as the suppressor click was presented at a level that was 15 dB lower than the level of the test click. A trade-off exists between the effectiveness of the suppressor (i.e., high level suppressor) and stimulus-synchronization of the CEOAE. The latter can not be omitted and thus the effectiveness of the suppressor needs to be kept low. It was estimated that the actual compression would span a range of about 5 dB when removing the suppressor-induced phase component, for the case when suppressor and test click were presented at 65 dB peSPL. Compression can thus be investigated correctly by the use of suppression in CEOAEs, only the effect

is small. This makes temporal suppression of CEOAEs unattractive as a method for investigating temporal changes in the compression of the CEOAE level-curve.

2.4.2 Temporal suppression of SOAEs and CEOAEs

Stimulus-synchronization worked very similar for the SOAE and CEOAE components. For the situation where both the suppressor and test click had the same level (Fig. 2.7), there were traces that were designated as not having a perfect synchronization with the click stimuli. For some of these traces, the short-latency CEOAE synchronized but the long-latency CEOAE did not (e.g. ICIs of 5, 6 and 7 ms in Fig. 2.7). When the level difference between the suppressor and test click increased, both the short- and long-latency CEOAE synchronized to the click stimuli (Fig. 2.8). This indicated that the SOAEs are more robust before synchronizing to the stimulus than the CEOAE components, but when they do, then stimulus-synchronization of SOAEs and CEOAEs to the click stimuli is identical. Temporal suppression related to compression in the auditory system affected SOAEs and CEOAEs thus identically (Fig. 2.9). This emphasizes their similar origins in the cochlea, as was earlier suggested by Kemp and Chum (1980).

2.4.3 CEOAE generator mechanisms

A static nonlinearity, based on the CEOAE level-curve, was found to model suppression where maximal suppression occurs for coinciding suppressor and test click, reasonably well (Kapadia and Lutman, 2000a; Harte et al., 2005). A time-dependent nonlinearity was required next to the static nonlinearity to account for situations where maximum suppression occurred at ICIs other than zero for leading suppressor clicks (Harte et al., 2005). Hine and Thornton (2002) speculated this time dependency to originate from complex displacement patterns along the basilar membrane that could create a nonlinear temporal overlap at the generation sites of the CEOAEs. They thereby implied that the CEOAE generator mechanism in itself remained unaffected by the time dependency, but that the input to the CEOAE generator was distorted in a nonlinear fashion by presenting the suppressor click close in time to the test click. Kapadia and Lutman (2000b) hypothesized that temporal suppression for leading sup-

pressors arose from a 'disturbance', related to the envelope of the displacement pattern along the basilar membrane in response to the suppressor stimulus. Their approach is similar to that of Hine and Thornton (2002), but they speculate the mechanisms of temporal suppression not only to be linked to the basilar membrane excitation pattern, but also to the CEOAE generator element itself. The "disturbance" would be able to offset the CEOAE generator element from its resting state, so that when a test click is presented, the actual operating point of the CEOAE generator element could be shifted (Kapadia and Lutman, 2000b). They suggested that a modification of the static nonlinear model presented in Kapadia and Lutman (2000a) could account for this offset. This offset could be realized by either changing the compression factor of the CEOAE level-curve according to the ICI or by temporally shifting the operating point of the stimulus on the static CEOAE level-curve (Kapadia and Lutman, 2000b). Lina-Granade and Collet (1995) assumed that the generator mechanisms for temporal suppression originated either from nonlinearities in the traveling wave overlap or from an adaptation of the outer hair cells. Adaptation of outer hair cells by a saturation of the gating channels (i.e., the K^+/Ca^{2+} transduction channel in the stereocilia of the hair cell can be saturated when deflecting the hair cell from its resting state (Hudspeth and Gillespie, 1994)) in response to the suppressor click could be the source responsible for the offset of CEOAE generator mechanisms during second click stimulation, as suggested by Kapadia and Lutman (2000b).

The key to favor one of the proposed hypotheses on the generator mechanism of temporal suppression lies in the augmentation that was observed in the presented results. Augmentation in different types of evoked emissions has been observed and discussed before, but the mechanisms behind it are still under debate. A study by Withnell and Yates (1998) showed that when pure tones of 70 dB SPL were presented together with electrical clicks in guinea pigs, enhancement of certain TEOAE components apical to the pure tone characteristic place was observed. They postulated that inter-modulation energy under the form of distortion products might suppress or augment certain TEOAE components and that the exact degree of suppression would depend on the summation of the inter-modulation products. Ren and Nuttall (1999) contested this hypothesis and proposed another, passive reflection theory. Augmentation of certain TEOAE components below the characteristic frequency of the pure tone

was explained by traveling-wave reflections from the characteristic frequency place of the pure tone rather than the characteristic place of the components within the click stimulus (Ren and Nuttall, 1999). Yates and Withnell (1999) observed no changes in the group delay of the TEOAE when adding the pure tone to the click and therefore rejected the passive reflection theory of Ren and Nuttall (1999). Kalluri and Shera (2007) have recently rejected the inter-modulation theory of Withnell and Yates (1998) with the proof that SFOAEs and CEOAEs share similar generator mechanisms at low to moderate stimulation levels. Electrically evoked otoacoustic emissions (EEOAE) in guinea pigs showed augmentation of the response when presenting acoustical pure tones in combination with the electrical stimulation (Mountain and Hubbard, 1989; Xue et al., 1993; Kirk and Yates, 1996, 1998). The idea that a small shift in the outer hair cell operating point could explain this augmentation of the EEOAEs (Kirk and Yates, 1998) seems to relate to the operating-point shift of the static compressive level-curve proposed by Kapadia and Lutman (2000b).

Although the above mentioned studies cannot directly be compared to one another because they adopt different research methods, they all study the evoked emission generator mechanisms. It was demonstrated here that augmentation in CEOAEs was linked to a cochlear effect, and this was the main reason for rejecting the hypotheses of Hine and Thornton (2002) and Lina-Granade and Collet (1995). The nonlinearities in the basilar membrane excitation patterns alone cannot account for all aspects of temporal suppression and thus the static nonlinear models of Kapadia and Lutman (2000a) and Harte et al. (2005) need modifications to account for augmentation. The excitation patterns might still play a role in the suppression process when discussing stimulus-level dependencies in suppression, but a change of the CEOAE generator mechanisms is needed to explain augmentation as well as suppression.

The approach of Kapadia and Lutman (2000b) and Lina-Granade and Collet (1995) to offset the static nonlinearity in the CEOAE generator mechanisms formed the basis of the phenomenological model presented here. Fig. 2.10 illustrates how temporal suppression and augmentation can conceptually be created with a CEOAE level-curve that is shifted depending on the ICI. The idea is that the system, in response to a single click, would shift the knee-point of the level-curve horizontally during a certain time period, before the level-curve would become static and the operation steady state.

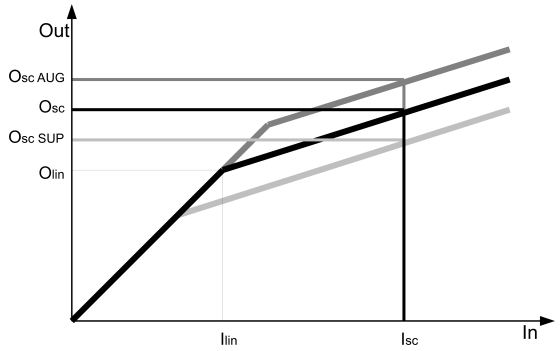


Figure 2.10: Illustration of how suppression and augmentation can conceptually be created with a compressive level-curve. O_{sc} is the level of a CEOAE to a single click (I_{sc}) and is referred to as the reference situation. The input range of the I_{sc} below I_{lin} is processed linearly whereas the input range above I_{lin} is compressed by a factor n, represented by the slope of the level-curve. Presenting a second click close in time to the first one could conceptually change the level-curve that is seen by the second click. Suppression of the output $O_{sc}SUP$ is obtained with a leftward shift of the knee-point of the level-curve (light gray); augmentation $O_{sc}AUG$ is created with a rightward shift of the knee-point (dark gray).

This time period would correspond to the duration of the temporal suppression effect. If a second click is presented during this time period, then the second click would see a modified level-curve. As a result, the output to the second click would have a level that is different from the output to the first click. The output to the first click is shown as O_{sc} in Fig. 2.10 and is referred to as the reference situation. The input level range with levels below I_{lin} is processed linearly and the input range above I_{lin} shows compression with a factor n, that represents the slope of the level-curve. Suppression is obtained by shifting the knee-point between linear and compressive behavior of the level-curve temporarily to the left (light gray level-curve in Fig. 2.10). In comparison to the reference situation, a smaller part of the dynamic range of I_{sc} is processed linearly and a larger part is processed compressively, leading to an output O_{scSUP} that is suppressed. Augmentation is found by shifting the knee-point of the level-curve in the rightward (dark gray level-curve in Fig. 2.10). As a consequence, a larger part of the dynamic range of I_{sc} is processed linearly leading to an output O_{scAUG} that is augmented in relation to O_{sc}. This phenomenological model can qualitatively account

for the data presented in this study when assuming that a, possibly subject-dependent, ICI-based shift of the CEOAE level-curve is physiologically possible.

2.5 Conclusion

1) Temporal suppression consists of two main components: suppression due to compression in the CEOAE generator mechanisms and suppression due to phase interactions. These phase interactions are caused by interactions between the ICI and the CEOAE components (method 1) or poor stimulus-synchronization of the CEOAE (method 2). Suppression in method 1 can never be interpreted in terms of compression because the ICI-related phase component can not be removed. Method 2 does allow for a comparison between the observed suppression and changes in compression of the CEOAE generator mechanisms, when assuring that the stimulus-synchronization of the CEOAE is present. This can be done by choosing a suppressor level that is lower than the test level, which leads to a reduced effectiveness of the suppressor click. Assuring good stimulus-synchronization conditions leads to a temporal suppression effect that has a small dynamic range. It is concluded that method 2 is capable of investigating compression by analyzing temporal suppression in CEOAEs, but is not recommended because of the small dynamic range.
2) Suppression related to compression can show both augmentation and suppression, depending on the subject under test. This leads to underlying CEOAE generator mechanisms that need to be capable of both compressing and expanding. This can be realized by an ICI-related shift of operation along the CEOAE level-curve when a second click is presented close in time to a test click. Further experimental evidence is needed to prove the existence of such time-dependent operating-point shifts.
3) The SSOAE and CEOAE components are similarly affected by temporal suppression, when there is stimulus-synchronization of the CEOAE. This provides evidence for stating that SOAEs and CEOAEs share the same underlying mechanisms.

3

Temporal Adaptation of the Click-Evoked Otoacoustic Emission Level-Curve [1]

The level of a click-evoked otoacoustic emission (CEOAE) depends on the level of the evoking click through the CEOAE level-curve. The CEOAE level grows linearly for clicks below 40-60 dB and saturates for higher input levels. This study shows that the CEOAE level-curve for a test click alters when a suppressor click is presented less than 8 ms before the test click. This dynamic (i.e., time-dependent) change of the level-curve is referred to as temporal adaptation of the CEOAE level-curve, and its strength is determined by the levels and the temporal separation of the two clicks. Temporal adaptation occurred for all four subjects tested, and for the short- (i.e., 6–18 ms) and long-latency (i.e., 24–36 ms) region of the CEOAE. This indicated that temporal adaptation affected synchronized spontaneous otoacoustic emissions (SSOAEs) and purely evoked CEOAE components similarly. Temporal adaptation was furthermore found to consist of a component common to all subjects, and a subject-dependent component that described the magnitude and time scale of the effect. Dynamic changes of the CEOAE level-curve were shown to describe dynamic properties of human cochlear processing, and suggest the existence of a short time-constant (0–10 ms) in the generator mechanisms of the CEOAE.

[1] This chapter formed the basis for Verhulst et al. (2011) and was presented on the 2009 Midwinter Research Meeting of the Association for Research in Otolaryngology in Baltimore, ML

3.1 Introduction

Click-evoked otoacoustic emissions (CEOAEs) can be measured in 98% of the normal-hearing population and are characterized by a broad frequency spectrum (Probst et al., 1991). The fine-structure of a CEOAE spectrum reflects subject-dependent patterns of basilar membrane (BM) impedance irregularities, which are considered to be reflection sources (Kemp and Martin, 1976; Shera and Guinan, 1999). When a forward traveling wave (e.g. caused by acoustical stimulation) encounters these irregularities, microscopic reflections can occur and sum coherently. These coherent reflections are believed to give rise to a reverse traveling wave, that can be measured in the ear canal as an OAE (Zweig and Shera, 1995). The forward traveling wave in response to a click excites nearly all frequencies along the basilar membrane, which leads to a CEOAE spectrum that contains reflections from various BM locations. The purely stimulus-evoked component of a CEOAE typically lasts 20–30 ms (Kemp, 1978; Probst et al., 1991), and forms the short-latency part of the CEOAE that is recordable in most normal-hearing subjects. Spontaneous otoacoustic emissions (SOAEs) can synchronize to the evoking click, and thereby create long-lasting synchronized SOAEs (SSOAEs; Wilson, 1980) with latencies up to 200 ms. SOAEs are thought to be a consequence of cochlear standing waves that arise in the cochlea even if no external stimulation is applied (Shera, 2003). SOAE and purely evoked CEOAE components coexist in long-lasting CEOAEs (i.e., duration > 30–40 ms), that reflect a short-latency CEOAE component dominated by cochlear reflections to the forward traveling wave, and a long-latency component attributed to amplitude stabilized cochlear standing waves.

The level of a CEOAE depends on the level of the evoking click, and this is reflected in the CEOAE level-curve (Kemp, 1978). The CEOAE level grows linearly for clicks below 40–60 dB and saturates for higher input levels (Kemp, 1978; Probst et al., 1991; Kalluri and Shera, 2007). The compressive behavior of the CEOAE level-curve is a consequence of compressive nonlinearities in the cochlea, making CEOAEs an interesting tool for investigating BM compression in humans. Dynamic, or time-varying, features of CEOAE and BM compression have been investigated in the past by means of temporal suppression in CEOAEs (Kemp and Chum, 1980; Tavartki-

ladze et al., 1994; Lina-Granade and Collet, 1995; Kevanishvili et al., 1996; Kapadia and Lutman, 2000a,b; Hine and Thornton, 2002; Verhulst et al., 2008). Temporal suppression is defined as the reduction in CEOAE level when a suppressor click is presented close in time (e.g., 0–10 ms) to a test click. The most extensive study on temporal suppression has been conducted by Kapadia and Lutman (2000a,b). They found that suppression, due to the presentation of a preceding click, was maximal for inter-click intervals (ICIs) of 2–4 ms and reduced drastically for ICIs larger than 12 ms. Furthermore, Kapadia and Lutman (2000a,b) found that suppression maxima moved to shorter ICIs when increasing the level difference between the two evoking clicks. Verhulst et al. (2008) showed that suppression levels and maxima were subject dependent and that suppression can also be observed for long-latency CEOAEs dominated by SSOAEs, while Kapadia and Lutman (2000a,b) only considered short-latency emissions.

In the past, two cochlear mechanisms have been linked to temporal suppression observed in CEOAEs: (i) temporal overlap of the BM excitation patterns of the two clicks (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Hine and Thornton, 2002); and (ii) a temporal change in the BM input/output function caused by the presentation of the suppressor click (Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Verhulst et al., 2008). Kapadia and Lutman (2000a) and Harte et al. (2005) have investigated the mechanisms underlying suppression by modeling temporal suppression with a static nonlinearity that simulated BM compression. These conceptual models could qualitatively describe decreased suppression for increased spacing between the clicks, but could not account for the exact time scale of suppression. The models of Kapadia and Lutman (2000a) and Harte et al. (2005) considered only one specific location along the cochlear partition. They thereby assumed that the CEOAE level-curve would be identical to the BM I/O function that describes the relation between stimulus level and BM velocity/displacement at that specific location (e.g. Robles et al., 1986). It is postulated here that the CEOAE level-curve and BM I/O functions cannot directly be compared to each other, but that the broadband CEOAE level-curve is a consequence of local BM I/O functions. This statement was supported by the observation that the knee-point of the CEOAE level-curve occurs at higher stimulus levels (40–60 dB) than for

the stimulus frequency otoacoustic emission (SFOAE) level-curve (20–30 dB; Kalluri and Shera, 2007). Changes in the CEOAE level-curve due to temporal suppression may therefore represent an across-frequency temporal effect, which cannot directly provide information about temporal processes at local sites on the BM. One way to avoid this across-frequency component in temporal suppression would be to look at similar dynamic suppression effects in SFOAEs or in distortion-product otoacoustic emissions (DPOAEs). However, the introduction of a suppressor tone close to or at the test frequency would make it difficult to visualize OAE suppression on a time scale of 0 to 10 ms, due to the available OAE extraction techniques. CEOAEs, and potentially transient-evoked otoacoustic emissions (TEOAEs; e.g. toneburst OAEs), are thus to date the only objective and non-invasive techniques to investigate short-time suppression effects in human cochlear processing.

This study considers temporal suppression to be the change from a known CEOAE level-curve to an *adapted* CEOAE level-curve that is obtained when a suppressor is presented 0–10 ms before the test click. Even though previous studies have demonstrated a relation between the amount of suppression and the level of the evoking click (Kemp and Chum, 1980; Kapadia and Lutman, 2000b), they did not explicitly address dynamic features of the CEOAE level-curve for specific time intervals between the clicks. The CEOAE level-curve reflects subject-dependent features such as compression threshold and slope, and those features are lost when analyzing suppression without considering the original CEOAE level-curve (Kemp and Chum, 1980; Kapadia and Lutman, 2000b). It is furthermore crucial, when adopting temporal suppression to investigate cochlear mechanisms, that the changes ascribed to temporal suppression stem from magnitude (i.e., not waveform) changes in the CEOAE (Verhulst et al., 2008). The studies of Tavartkiladze et al. (1994), Kevanishvili et al. (1996) and Hine and Thornton (2002) used temporal suppression as a measure for both waveform and magnitude changes of the CEOAE when a suppressor was presented close in time. However, because waveform changes do not arise from compression changes in the underlying mechanisms, the results from these studies cannot be used to investigate dynamic changes in cochlear compression (Verhulst et al., 2008).

CEOAE level-curves can be obtained by plotting the rms level of a specific time-window of the CEOAE as a function of stimulus level. The left panel of Fig. 3.1 shows

a conceptual CEOAE level-curve that shows compressive growth of the CEOAE in response to a linearly increasing stimulus level, after a certain threshold level was reached. Kemp and Chum (1980), Kapadia and Lutman (2000b) and Verhulst et al.

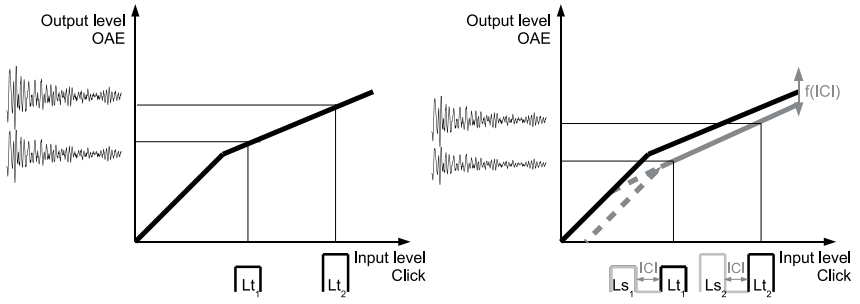


Figure 3.1: Concept of the CEOAE level-curve (left). The hypothesis of this study (right) stating that the CEOAE level-curve to a test click shifts horizontally when a suppressor is presented at a time interval of 0–10 ms before a test click. For the same L_t as in the unsuppressed condition (black curve), it is expected that the CEOAE to the L_t in the suppressed condition (i.e., L_t at an ICI after L_s; red curve) shows a reduced level. Two different configurations for the knee-point in the suppressed level-curve are indicated. Note that L_s was plotted in gray and not in black, because the suppressed condition describes the CEOAE level to the test click after suppressor presentation, and not the CEOAE level to the combined stimulus of suppressor and test click.

(2008) showed that the level of a CEOAE typically reduces in the presence of a preceding suppressor, and therefore Kapadia and Lutman (2000b) designated this CEOAE as being suppressed. The *suppressed* condition is illustrated in Fig. 3.1, where the CEOAE to the test click, presented at a certain ICI after a suppressor (L_{t1}, right panel), was reduced in level compared to the *unsuppressed* CEOAE to the test click alone (L_{t1}, left panel). It is hypothesized here that, if the unsuppressed and suppressed CEOAE similarly depend on stimulus level, the CEOAE level-curve shifts horizontally when a suppressor is presented at a time interval of 0–10 ms before the test click. Based on suppression results from Kemp and Chum (1980), Kapadia and Lutman (2000b) and Verhulst et al. (2008), it is furthermore expected that the degree of shifting depends on the time and level difference between the two clicks. In order to fully describe the influence of a preceding suppressor on the CEOAE level-curve, it is furthermore investigated whether the knee-point (i.e., the transition region between

the linear and compressive region) of the CEOAE level-curve remains at the same threshold level for the suppressed CEOAE.

3.2 Materials and Methods

3.2.1 Deriving the suppressed CEOAE level-curve

The suppressed CEOAE cannot be measured directly, but needs to be derived from the response recorded to the combined suppressor and test click stimulus. The typical length of a CEOAE (20–30 ms for the short-latency and up to 200 ms for the long-latency components) causes the CEOAE to the suppressor click to overlap in time with the CEOAE to the test click. The response to the combined double-click stimulus thus consists of a CEOAE to the suppressor click, a CEOAE to the test click and a nonlinear component that depends on the ICI. The interaction between these three components is usually observed as a combined response that is smaller than twice the response to a test click alone (Kemp and Chum, 1980; Kapadia and Lutman, 2000b; Verhulst et al., 2008). Kemp and Chum (1980) developed a technique to remove the CEOAE component from the suppressor click, while keeping the CEOAE component to the test click and the nonlinear component due to the ICI. This technique, as adapted by Kapadia and Lutman (2000b), was illustrated in Fig. 3.2.

Two double-click stimuli were presented for every ICI and level configuration tested: DC (i.e., double-click stimulus) and DCI (i.e., double-click stimulus with an inverted test click). DC consisted of a condensation suppressor and test click; DCI consisted of a condensation suppressor click and a rarefaction test click. The derived suppressed response, DS, was obtained by subtracting DCI from DC and by halving the resulting response. In DS, the linear response to the suppressor was removed while the nonlinear component, attributed to the ICI, remained. Suppression was defined as the level difference between DS and US (i.e., unsuppressed) response, with the latter recorded to the test click alone. Applying this measurement paradigm for a specific ICI and click level configuration leads to one estimated point on two different CEOAE level-curves: one on the unsuppressed CEOAE level-curve and one on the suppressed CEOAE level-curve (see Fig. 3.2). The adopted measurement paradigm

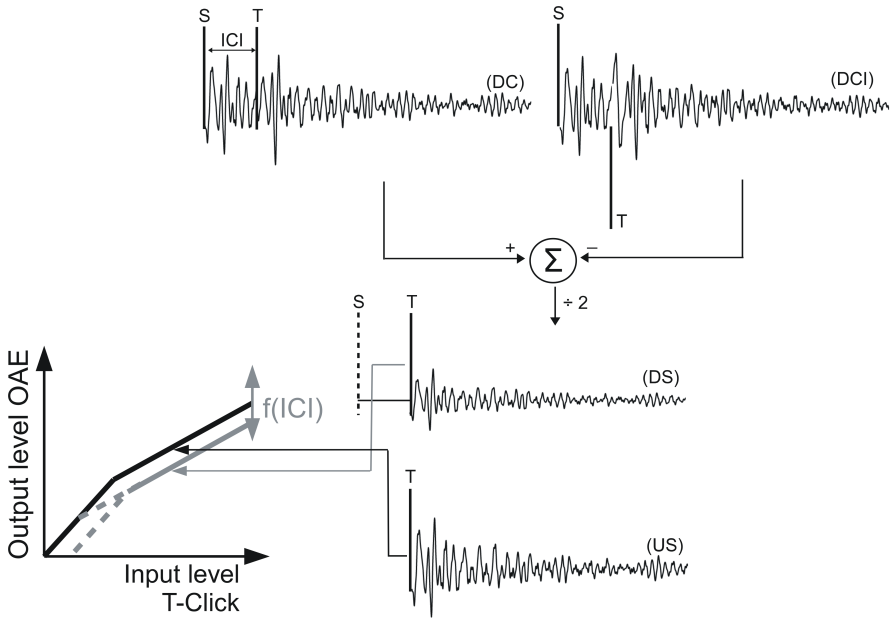


Figure 3.2: The measurement paradigm adopted to obtain the unsuppressed and suppressed CEOAE. DC consisted of a condensation suppressor (S) and test click (T), separated by a certain ICI; DCI consisted of a condensation S-click and a rarefaction T-click. The derived suppressed response, DS, was obtained by subtracting DCI from DC and by halving the resulting response. US was recorded to the T-click alone. Suppression was calculated as the level difference between DS and US. Applying this measurement paradigm for a specific ICI and level configuration leads to one estimated point on two different CEOAE level-curves: one on the unsuppressed and one on the suppressed level-curve.

aimed at investigating cochlear compression mechanisms. To rule out measurement biases that were not due to compression, a couple of preliminary experiments were conducted and presented in Appendix A.

The stimuli, consisting of 80 μs long clicks, were presented to the subjects in an interleaved manner: for every ICI and stimulus level configuration, an epoch consisting of US, DC and DCI was presented before the next epoch consisting of US, DC and

DCI was presented. Each epoch consisted of a 50-ms[1] long US, DC and DCI window (comprising stimuli and CEOAE), and 2000 epochs were recorded for every ICI and stimulus level configuration.

Two experiments were conducted in this study. In the first experiment, the time scale of suppression for different suppressor-click levels (L_s) was investigated. Suppression was investigated for a test-click level (L_t) of 65 dB peSPL and an L_s of 59, 62 and 65 dB peSPL. ICIs between 0 and 8 ms were considered to show the accordance of our results with previous studies (Kemp and Chum, 1980; Kapadia and Lutman, 2000b; Verhulst et al., 2008). The experimental conditions are summarized in Table 3.1. In the second experiment, unsuppressed and suppressed CEOAE level-curves were derived for different ICIs. L_t varied between 30 and 85 dB peSPL for the unsuppressed condition, and between 45 and 75 dB peSPL for the suppressed condition. Suppressed level-curves were derived for ICIs of 1, 2 and 3 ms, and for L_s = L_t and 3, 6, 9 dB lower than L_t. Table 3.2 details the conditions for experiment 2.

Table 3.1: Levels of the suppressor (L_s) and test clicks (L_t) that were tested in experiment 1 for different ICI conditions. The experiment was split up in three sessions (M#) and performed on all four subjects in this study.

L_t	L_s [dB peSPL]	M#	ICI [ms]
65	65	1	0, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 7, 8
65	62	2	0, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 7, 8
65	59	3	0, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 7, 8

3.2.2 Subjects

Two subjects with short-latency CEOAEs and two subjects with long-latency (duration > 30–40 ms) CEOAEs participated in this study. The subjects with long-latency CEOAEs were included to enable an investigation of suppression in the evoked and

[1] There was a possibility that the SSOAEs of one recording window (50 ms) would continue to ring into the next recording window. This could potentially complicate the generation of the SSOAE in the next recording window. The influence of existing SSOAEs on the generation of an SSOAE to a click was investigated for one subject that had long-latency SSOAEs. Ten window durations between 28.3 and 50 ms were tested, and there was no apparent influence of the window duration onto the recorded CEOAE. This implied that the relative phase of the SSOAE at the start of a click stimulus did not influence the generation of the SSOAE to this click stimulus.

Table 3.2: Levels of the suppressor (L_s) and test clicks (L_t) that were used to derive the unsuppressed and suppressed CEOAE level-curves for different ICI conditions. The conditions marked with – indicate levels that were used for the unsuppressed US condition. The ICIs marked with * were measured for subject 1,3 and 4 and the ICIs marked with ** were measured for subject 1 and 2. All other ICI and level configurations were measured for all four subjects. The duration of the measurement sessions (M#) was around one hour.

L_t	L_s [dB peSPL]	ICI [ms]						
	M#	1	2	3	4	5	6	7
30	–	–						
35	–	–						
40	–	–						
45	–, 45, 42, 39	–	2			1*, 3**		
50	–, 50, 47, 44	–	2			1*, 3**		
55	–, 55, 52, 49	–	2			1*, 3**		
60	–, 60, 57, 54, 51	–		2			1*, 3**	
65	–, 65, 62, 59, 56	–			2			1*, 3**
70	–, 70, 67, 64, 61	–		2			1*, 3**	
75	–, 75, 72, 69, 66	–			2			1*, 3**

SSOAE components of the OAE. All subjects had normal hearing abilities (i.e., thresholds below 15 dB HL across the frequencies considered in a standard audiogram). During the recordings, subjects were asked to lie down in a soundproof booth with the probe in the left ear-canal. The probe was not removed from the ear-canal during the one-hour long measurement sessions, and subjects were asked to relax, refrain from moving and to stay awake.

3.2.3 Apparatus

The stimuli were generated in Matlab and sent to a soundcard (RME FireFace 800 A/D-D/A converter) via the open source software pa-wavplay[2]. The signal levels were calibrated to peak-equivalent sound pressure level (peSPL) in a BK-2012 ear-canal coupler, attached to a BK-4157 artificial ear calibrator. Stimulus levels were controlled via a DT-PA5 programmable attenuator. The stimuli were presented to the

[2] pa-wavplay contains a set of dll's allowing simultaneous play-rec for multiple channels. The code was written by Matt Frear (Univ. of Western Sydney) and uses the open source portaudio API. The open source code of pa-wavplay can be found on http://sourceforge.net/projects/pa-wavplay (last visited: 26^{th} of October 2009)

test subjects via two ER-2 earphones such that condensation clicks were presented over the first channel and rarefaction clicks over the second. Recordings were made using an ER-10B+ low noise microphone, and were bandpass filtered between 0.6 and 5 kHz (analog Rockland 852 HI/LO filter). The recorded analog signals were converted and stored digitally.

3.2.4 Post-processing

For every ICI and stimulus-level condition measured in this study, 2000 epochs of a 150-ms long US–DC–DCI recording window were obtained. In the post-processing stage, the DS response was derived by subtracting the corresponding DC and DCI epochs and by halving the resulting waveforms, after which the 2000 DS responses were averaged. An averaged US reference was also obtained for every ICI and stimulus-level configuration. To increase the signal-to-noise ratio of the averaged US and DS responses, an artifact-rejection template was applied to remove $\sim$10% of the noisiest epochs before averaging. The early-latency linear response, associated with the ear-canal and middle-ear transfer function, was removed from the US and DS responses to derive the CEOAE[3]. Further analysis was carried out on the CEOAE, set to start $\tau_s = 6$ ms after test click onset for all subjects.

Root-mean-square (rms) levels of the unsuppressed and suppressed CEOAE were calculated for several time frames of the CEOAE between $\tau_s = 6$ and 36 ms after test click onset with:

$$L(t_p) = 20\log_{10}\left[\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2\right]^{1/2}, \tag{3.1}$$

where N depicts the number of samples in the analysis window. x_i represents the i^{th}

[3] To determine this separation point from the early-latency response respective to the CEOAE, two US responses were measured at 65 and 75dB peSPL. Both responses were normalized and plotted on top of each other. The linear early-latency components sat atop, and the CEOAE components diverged due to CEOAE compression. This technique has the advantage of giving a time point to separate the CEOAE-dominated response from the early-latency response for a given subject, without actually affecting the CEOAE time series. The linear response in the click recordings decreased sufficiently between $\tau_s = 5$ and 6 ms after click onset for all subjects.

sample of the CEOAE and μ_x stands for the mean of the CEOAE within the considered window. A value for L_{rms} was found for every time frame t_p of the CEOAE. rms Levels were calculated between $\tau_s = 6$ and 18 ms for the short-latency CEOAE, and between $\tau_s = 24$ and 36 ms for the long-latency CEOAE[4].

Suppression was calculated as:

$$\text{Suppression}(t_p) = L_{US}(t_p) - L_{DS}(t_p) \tag{3.2}$$

where $L_{US}(t_p)$ and $L_{DS}(t_p)$ were the rms levels of US and DS in the corresponding time frame. L_{US} and L_{DS} furthermore represented data points on the unsuppressed and suppressed CEOAE level-curve respectively, for a specific ICI and stimulus level configuration. Suppression was thus a measure for the vertical deviation from the unsuppressed level-curve.

Standard deviations, representing the variability of the suppression measure within one measurement session, were calculated for all suppression levels. Five L_{US} and L_{DS} were calculated for every 'sub' average of 360 DS and US responses. When cross-subtracting the 5 L_{DS} from the 5 L_{US}, 25 suppression levels were obtained. The standard deviation of these 25 suppression levels represented the variability of suppression during a specific measurement session, and were indicated for the calculated suppression levels. The repeatability of the suppression measure across measurement sessions was furthermore tested by recording the same condition (i.e., ICI = 2 ms, L_s = L_t = 65 dB peSPL) in every measurement session. The repeatability for suppression was found to be 0.57, 0.55, 1 and 0.6 dB for the short-latency CEOAE of subjects #1 to #4, and 1.5 and 1.7 dB for the long-latency CEOAE of subjects #1 and #2.

[4] Note that the short- and long-latency regions were chosen to represent windows where the evoked (short) and SSOAE (long) components were expected to be more pronounced. The choice for these windows was made arbitrarily and not based on a specific component separation technique. Withnell et al. (2008) and Goodman et al. (2009) on the other hand, have defined *early* and *late*-latency regions of the CEOAE depending on the underlying cochlear mechanisms. Properties of early-latency (< 5–6 ms) CEOAEs were shown to be indicative of inter-modulation distortion mechanisms, whereas late-latency CEOAEs (> 6 ms) agreed with the coherent reflection theory. (Shera and Guinan, 1999; Kalluri and Shera, 2007; Withnell et al., 2008; Goodman et al., 2009). The short- and long-latency CEOAE in this study were situated in the late-latency region of Withnell et al. (2008) and Goodman et al. (2009)

3.3 Results

3.3.1 Time scale of temporal suppression

Fig. 3.3 shows temporal suppression in function of ICI for the short- and long-latency CEOAE. As L_t was 65 dB peSPL, Fig. 3.3 reflects the deviation of the 65 dB peSPL

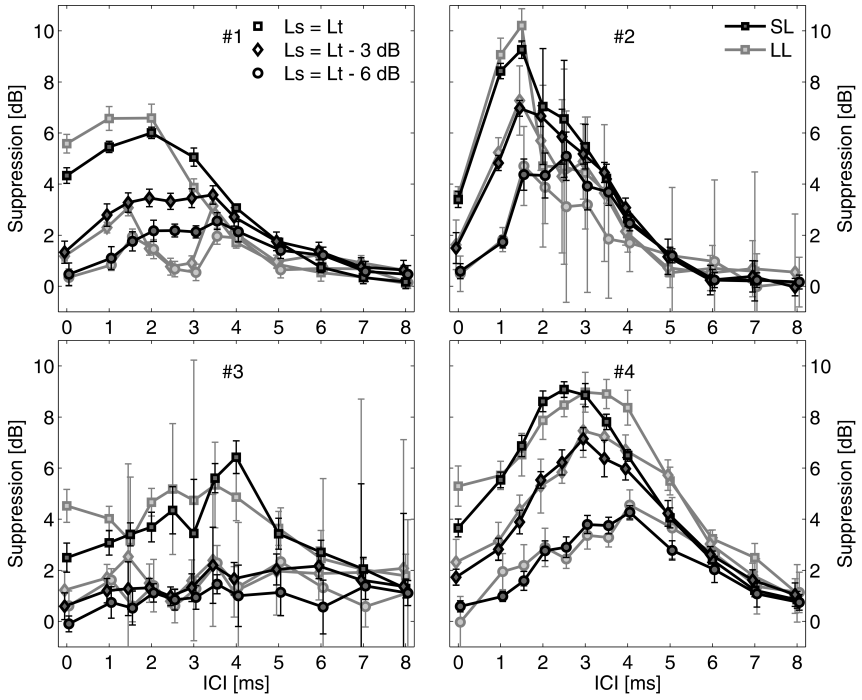


Figure 3.3: The deviation (i.e., suppression) of the 65 dB point on the unsuppressed CEOAE level-curve as a function of ICI and suppressor level. Suppression was calculated for the short-latency (6–18 ms; black curves) and the long-latency (24-36 ms; gray curves) CEOAE of the four subjects in this study. L_t was 65 dB peSPL and L_s was either 65, 62 or 59 dB peSPL. The ICIs tested were: 0, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 7 and 8 ms. The standard deviations reflect the variability of suppression during the measurement session. Note that the suppression curves for different L_s conditions were offset horizontally by 0.1 ms to aid interpretation.

point on the unsuppressed CEOAE level-curve. All four subjects showed suppression levels that increased with increasing ICI until maxima were reached for ICIs between

1 and 4 ms. Suppression reduced and decayed down to zero for ICIs between 5 and 8 ms. Maximal suppression values ranged between 6 dB for subject #3 and 10 dB for subject #2. Suppression decreased with decreasing L_s, and for subjects #1, #2 and #4 the peak of the suppression pattern moved to higher ICIs when L_s decreased. For the shorter ICIs (< 3–4 ms), suppression decreased by 2–3 dB per 3 dB L_s reduction. For the longer ICIs (> 4 ms), suppression decreased by a lesser amount with decreasing L_s.

Temporal suppression for the short-latency CEOAE corresponded well to the mean suppression data of Lina-Granade and Collet (1995), Kapadia and Lutman (2000b), and Verhulst et al. (2008). Lina-Granade and Collet (1995) found average suppression to vary between 0 and 10 ms with a maximum of 5 dB for an ICI of 2 ms for L_s = L_t = 65 dB peSPL.Kapadia and Lutman (2000b) found similar results for L_t = L_s = 60 dB peSPL, and furthermore showed that the suppression maxima moved to an ICI of 3 ms when L_s = L_t - 10 dB. The correspondence between the short-latency data in Fig. 3.3 and the mean suppression data of Verhulst et al. (2008) was good, but differences were observed for individual suppression patterns. Verhulst et al. (2008) showed augmentation (i.e. negative suppression) for several ICIs in individual suppression patterns, whereas the data in this study did not.

Temporal suppression in Fig. 3.3 was also observed for the long-latency (24–36 ms) CEOAE, which was in agreement with the long-latency suppression observed in Verhulst et al. (2008). There were a couple of differences between suppression for the short and long-latency CEOAE in Fig. 3.3. Firstly, the suppression patterns for subject #1 and #2 moved toward smaller ICIs for the long-latency region of the CEOAE. Conversely, the suppression pattern of subject #4 moved by 1 ms to higher ICIs from the short- to long-latency CEOAE. Secondly, the sharpness of the suppression patterns was investigated by estimating the 3-dB bandwidth of the suppression patterns for L_t = L_s = 65 dB peSPL in Fig. 3.3. The width of the suppression patterns for the short-latency CEOAE was 5, 2.8, 4.5 and 2.8 ms for subjects #1 to #4. These values changed to 3, 1.5 and 4 ms for the long-latency CEOAE of subjects #1, #2 and #4 respectively. Lastly, the time scale of temporal suppression, which was determined by the ICI where suppression for L_s = L_t = 65 dB peSPL reduced to 1 dB, changed with latency. The time scale of suppression for the short-latency CEOAE was 6, 5, 8 and 7

ms for subjects #1 to #4. These values changed to 5, 5, 8 and 8 ms for the long-latency CEOAE respectively.

3.3.2 Relation between temporal suppression and CEOAE spectra

It was investigated whether there was a relation between suppression in Fig. 3.3 and the frequency content of the CEOAE. Fig. 3.4 shows CEOAE spectra for the short- (6–18 ms) and long-latency (24–36 ms) region for subjects #1 to #4. Because the

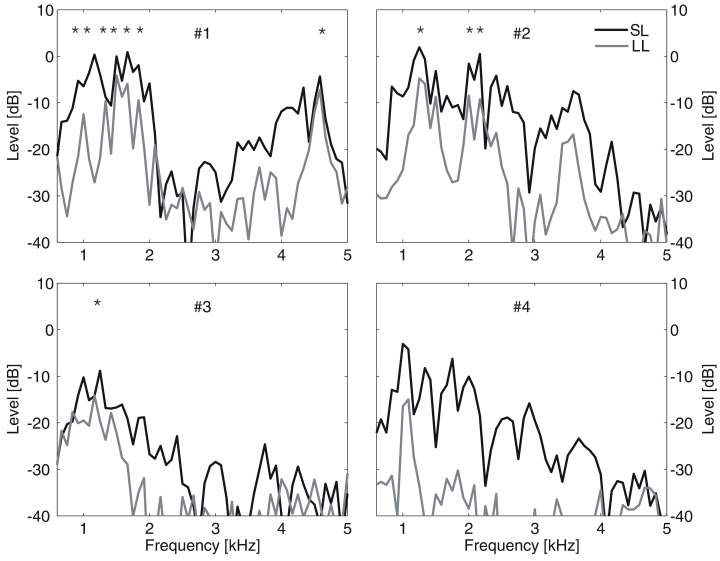


Figure 3.4: CEOAE frequency spectra of the four subjects subjects in this study. Spectra were plotted for the short- (6–18 ms) and long-latency (24-36 ms) regions of the CEOAE for each subject. The spectra corresponded to the calculated suppression levels in Fig. 3.3, 3.5 and 3.6. Strong SSOAE components were determined through a 130-ms-long CEOAE recording. Frequency components that were detectable for latencies longer than 40 ms were designated as strong SSOAE components, and were indicated on this figure by a *.

fine-structure of CEOAE spectra reflects subject-dependent cochlear reflection sites (Shera and Guinan, 1999), they provide information about which cochlear regions contributed to suppression. Subjects #2 and #4 in Fig. 3.4 showed broadband spectra with frequency components (i.e., above the -30 dB noise floor) in the 0.6–4.5 kHz

and 0.6–4 kHz region of the BM respectively. Subject #1 showed contributions to the CEOAE from the 0.6–2 kHz and 3.5–4.5 kHz region. Subject #3 showed a narrowband spectrum with frequency components that originated from the 1–2 kHz region on the BM. From the spectra in Fig. 3.4, it is furthermore observed that the fine-structure of the spectra changed over time. Subjects #1 and #2 represent the first type of subjects with strong SOAEs that can be observed in Fig. 3.4 as distinct frequency components (i.e., SSOAEs) that were present in the long-latency CEOAE. Subjects #3 and #4 represented the second type of subjects that had short-latency CEOAEs, and little or no SOAEs. From a 130-ms-long CEOAE recording of subjects #3 and #4, it was found that these subjects did not contain frequency components that lasted longer than 40 ms. The spectral components observed in the long-latency spectra of subjects #3 and #4, represent either evoked components that had not disappeared by the start of the long-latency window (i.e., 24 ms), or short-duration SSOAE components. Considering two types of subjects in this study allowed for an investigation of the suppression mechanisms underlying SOAEs (i.e., long-latency region of subjects #1 and #2) and purely evoked CEOAEs (i.e., short-latency region of subjects #3 and #4).

Maximal suppression levels were not related to the existence of SOAEs, as subjects #2 and #4 showed the largest amounts of suppression (10 and 9 dB respectively). Maximum suppression levels were furthermore not related to the energy in the CEOAE. Subjects #2 and #4 showed the largest amounts of suppression, and had broadband short-latency spectra with rms levels of 11.1 and 5.5 dB respectively. Subject #1 on the other hand, showed maximal suppression of 5.8 dB for a corresponding rms level of 9.43 dB. If suppression were related to the energy in the CEOAE, subject #1 would have larger suppression levels than subject #4. Furthermore, temporal suppression levels did not seem directly related to the frequency content of the emission. The maximal suppression level for subject #4 did not change from the short- to long-latency CEOAE. At the same time, the spectrum became narrowband, and the rms level of the CEOAE reduced by nearly 12 dB.

The sharpness of the suppression patterns were not linked to the presence of SOAEs, as the patterns of subject #2 and #4 were sharper than for subject #1 and #3. Again, a direct relation between the sharpness of the pattern and the frequency content of the emissions was not apparent. The only pronounced difference between

the spectra of subjects #1 and #3 and subjects #2 and #4, was that subjects #2 and #4 had more pronounced mid-frequency (2–4 kHz) components than subjects #1 and #3.

The time scale of suppression did appear to be related to the frequency content of the CEOAE. For the short-latency CEOAE, subjects #1 and #2 had the shortest time scales of temporal suppression and the most pronounced high-frequency content (> 4 kHz) in their spectra. The spectrum of subject #3 showed the least amount of high frequency components and showed the longest time scale for temporal suppression. For the long-latency CEOAE, the time scale of subject #1 changed from 6 ms to 5 ms, which corresponded to a reduction in spectral energy that was pronounced in the low (0.6-1.5 kHz) and mid (3–4.5 kHz) frequency region. The time scale of suppression for subject #4 changed from 7 to 8 ms, and showed a reduction from a broadband (0.6–4 kHz) to a narrow band (1–1.2 kHz) spectrum.

3.3.3 Temporal adaptation of the CEOAE level-curve

Fig. 3.5 shows the unsuppressed and suppressed level-curves for the short- and long-latency CEOAE of subjects #1 to #4. The suppressed level curves were derived for ICIs of 1, 2 and 3 ms and for L_s = L_t. The unsuppressed level-curves had levels between -15 dB and 20 dB for the short-latency (6–18 ms) CEOAE, and between -10 and 6 dB for the long-latency (24–36 ms) CEOAE. The long-latency level-curves for subjects #3 and #4 disappeared in the noise floor of the measurement. The unsuppressed, short-latency level-curves showed linear growth below the knee-point of the CEOAE level-curve, with slopes of 0.96, 0.89, 0.82 and 0.86 dB/dB for subjects #1 to #4. The slopes were calculated as the ratio between CEOAE level and L_t, for L_t from 40 to 50 dB peSPL. The knee-point between linear and compressive growth of the unsuppressed CEOAE level-curve in Fig. 3.3 was determined as the L_t that corresponded to the crossing of two straight lines that were fit to the linearly and compressively growing part of the level-curve. For the short-latency CEOAE, the knee-points were situated at stimulus thresholds of 55, 51, 62 and 56 dB peSPL for subjects #1 to #4. The knee-points of the long-latency level-curves of subjects #1 and #2 remained at the same stimulus threshold.

Above the knee-point of the level-curve, all subjects showed compressive growth with slopes of 0.27, 0.19, 0.51 and 0.35 dB/dB for subjects #1 to #4. The slopes were

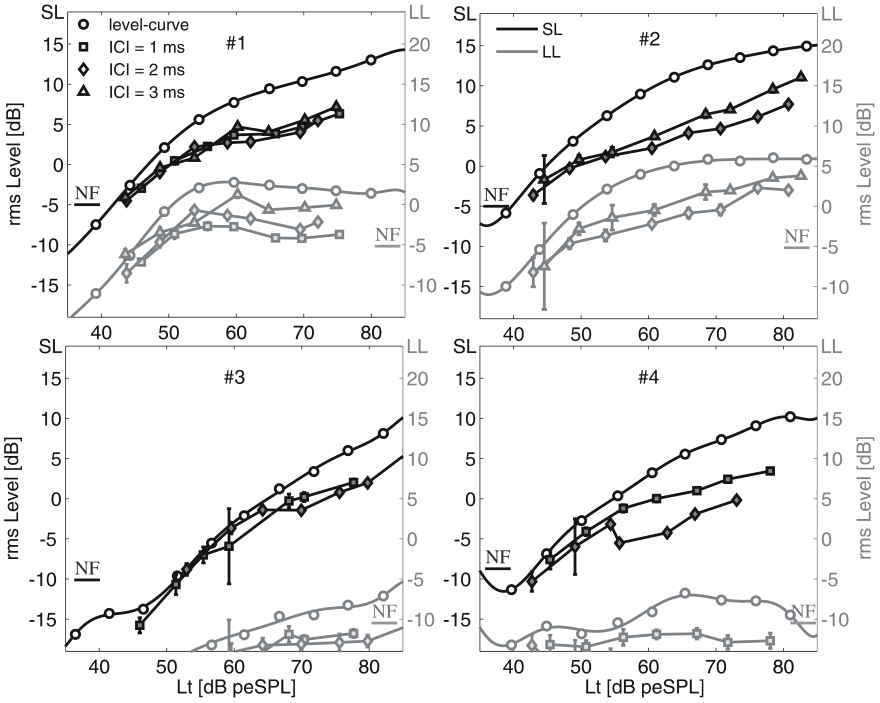


Figure 3.5: Derived CEOAE level-curves for the unsuppressed (circles) and suppressed (squares, diamonds, triangles) conditions for ICIs of 1, 2 and 3 ms and for $L_s = L_t$. The rms levels were plotted for the short-latency (SL; black curves) and the long-latency (LL; gray curves) CEOAE. All rms levels were plotted relative to the peak-to-peak level of the test click measured at the ear drum of the subject. Note that the vertical axis for the long-latency results was offset by 5 dB compared to the short-latency axis. The level-curves for the unsuppressed condition were obtained by fitting a ninth order polynomial to the data points measured on the level-curve. The suppressed levels were plotted as the deviation (i.e., suppression) from the corresponding level on the unsuppressed level-curve. The noise floor, determined as the threshold where the signal-to-noise ratio was less than 6 dB, was indicated with the 'NF' symbol for all subjects.

calculated as the ratio between the CEOAE level and L_t, for L_t from 60 to 80 dB peSPL. The slopes of the compressive regions of the long-latency level-curves changed to more saturated values of -0.07 and 0.04 dB/dB for subjects #1 and #2 respectively. The slopes of the unsuppressed level curves in Fig. 3.5 corresponded well to the slopes found in Kemp and Chum (1980): 0.45 dB/dB, Probst et al. (1986): 0.38 dB/dB, Prieve and Falter (1995): 0.35 dB/dB and Prieve et al. (1997): 0.32 dB/dB. The shallower

slopes found for the short-latency level-curves of subjects #1 and #2, compared to the slopes of subjects #3 and #4, agree well with the findings of Probst et al. (1986), Prieve and Falter (1995) and Prieve et al. (1997). They found that the CEOAE level-curves for subjects with SOAEs showed more compression than for subjects without, with mean slopes of 0.25, 0.28 and 0.33 dB/dB respectively. The short-latency level-curves for click levels between 40 and 80 dB peSPL ranged between -5 and 15 dB for subjects #1 and #2, and between -10 and 10 dB for subjects #3 and #4 in Fig. 3.5. Comparably, the levels of the short-latency level-curve were found to range between -3–7 dB and -5–9 dB for subjects without SOAEs, and between -2–13 and -5–12 for subjects with SOAEs in Prieve and Falter (1995) and Prieve et al. (1997) respectively.

All subjects in Fig. 3.5 showed suppressed level-curves that were different from the unsuppressed level-curves. The knee-points of the suppressed level-curves remained nearly at the same stimulus threshold for all suppressed conditions and latencies. The suppressed level-curves were observed as being vertically offset in the compressive region of the level-curve. The slopes of the compressive region of the suppressed and unsuppressed level-curves of subject #1 and #3 remained nearly the same, whereas the slope for subject #2 increased to 0.24 and 0.34 dB/dB for ICIs of 2 and 3 ms. Conversely, subject #3 showed shallower slopes with values of 0.21 (ICI = 1 ms) and 0.29 (ICI = 2 ms) for the compressive region of the suppressed versus unsuppressed CEOAE level-curve. The deviation of the suppressed CEOAE level-curve from the unsuppressed CEOAE level-curve was larger in the compressive region than in the linear region. Temporal suppression was either non-existing (subject #3) or gradually increasing with increasing stimulus level (subjects #1, #2 and #4) for the linear region of the CEOAE level-curve. Lastly, subjects #1 and #2 showed suppressed level-curves for the long-latency CEOAE that showed similar behavior in terms of slope change and knee-point location than the suppressed level-curves for the short-latency CEOAE.

The suppressed level-curves for different ICIs were not significantly different in the linear region, but were more pronounced in the compressive region. The time scale of temporal suppression was investigated in Fig. 3.3, and the results from experiment 1 for L_t = L_s = 65 dB peSPL corresponded well to the results found in Fig. 3.5 for L_t = 65 dB. Even though the two experiments were conducted on different days, the

repeatability of the temporal suppression measure was good. Thus, the time scale of suppression in Fig. 3.3 and the suppressed level-curves of Fig. 3.5 were complementary in describing the time scale and level dependence temporal suppression.

3.3.4 Level dependence of temporal adaptation

Fig. 3.6 shows the influence of suppressor level on temporal adaptation of the CEOAE level-curve for subject #1. It was observed that reducing L_s in steps of 3 dB led to an offset in the compressive region of the suppressed short-latency level-curve. For an ICI of 1 ms, this reduction was 6 dB when L_s = L_t, 4 dB when L_s = L_t - 3 dB, 2 dB when L_s = L_t - 6 dB, and 0 dB when L_s = L_t - 9 dB. The compressive regions of the level-curves for the ICIs of 2 and 3 ms showed similar reductions of 2 dB per L_s decrease of 3 dB. The slopes of the compressive region of the level-curve did not alter when reducing L_s. Though not plotted on Fig. 3.6, temporal suppression disappeared for all ICIs when L_s = L_t - 12 dB. Temporal suppression in the linear region of the CEOAE level-curve did not decrease much with decreasing L_s, but the linear region extended to higher stimulus levels. This was accompanied by a knee-point shift from 55 to 59 dB peSPL when decreasing the suppressor level by 6 dB. The level dependency of temporal suppression for L_t = 65 dB peSPL in Fig. 3.6 was in good agreement with the level dependency found in Fig. 3.3 for subject #1.

Subjects #3 and #4 (not shown) had suppression reductions of 2 dB per 3 dB reduction in L_s for the compressive region of the CEOAE level-curve. They furthermore showed knee-point shifts from 62 to 64 dB peSPL and from 58 to 60 dB peSPL when L_s reduced by 6 dB. Subject #2 (not shown) had a lesser reduction of suppression with decreasing suppressor level (1 dB per 3 dB), and a knee-point shift from 51 to 52 dB peSPL. Comparably, Kapadia and Lutman (2000b) found an average reduction in suppression of 4.8 dB per 10 dB suppressor for an L_t of 50 and 60 dB, and a reduction of 2 dB per 10 dB for an L_t of 70 dB.

3.3.5 Waveform dependence of temporal suppression

So far, temporal suppression and CEOAE level-curves were derived for short- and long-latency windows of the CEOAE. Suppression and temporal adaptation differ-

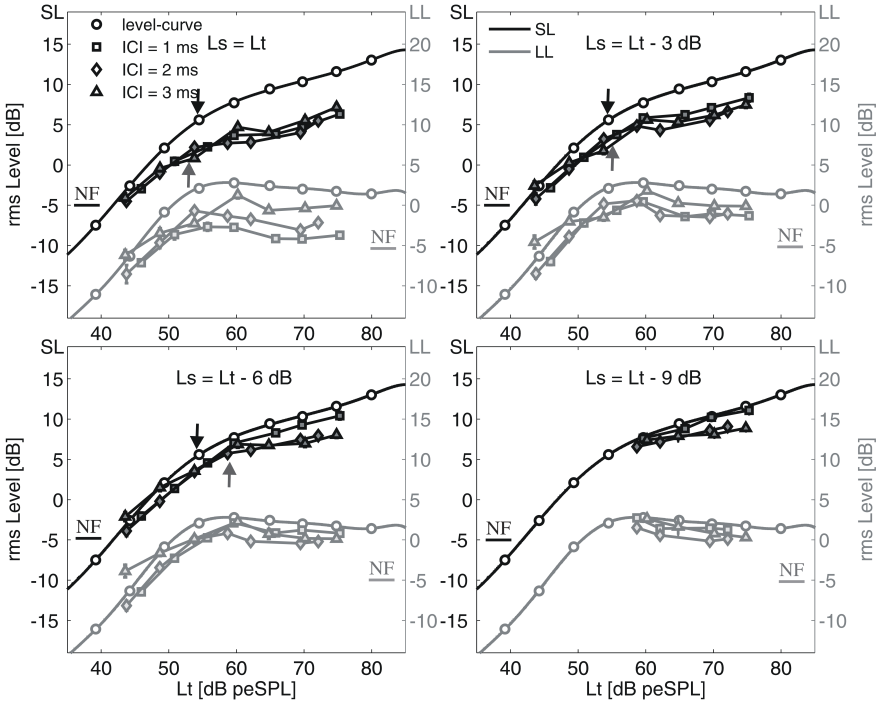


Figure 3.6: Derived CEOAE level-curves for the unsuppressed (circles) and suppressed (squares, diamonds, triangles) conditions for ICIs of 1, 2 and 3 ms and for four stimulus level conditions: $L_s = L_t$, or was 3, 6 or 9 dB lower than L_t. The rms levels were plotted for the short-latency (SL; black curves) and the long-latency (LL; gray curves) CEOAE of subject #1. All rms levels were plotted relative to the peak-to-peak level of the test click measured at the ear drum of the subject. Note that the vertical axis for the long-latency results was offset by 5 dB compared to the short-latency axis. The level-curves for the unsuppressed condition were obtained by fitting a ninth order polynomial to the data points measured on the level-curve. The suppressed levels were plotted as the deviation (i.e., suppression) from the corresponding level on the unsuppressed level-curve. The black and dark gray arrowheads indicate the knee-point of the unsuppressed and suppressed (ICI = 1 ms) level-curve respectively. The noise floor, determined as the threshold where the signal-to-noise ratio was less than 6 dB, was indicated with the 'NF' symbol for all subjects.

ences were pointed out for these latency regions in Fig. 3.3, 3.5 and 3.6, but it was not verified whether suppression was waveform dependent. This led to an investigation of instantaneous suppression across waveform latency in Fig. 3.8 and 3.9 for subjects #1 and #4. Instantaneous suppression (panel c) was defined as the difference between the envelope of the DS and US waveform for a specific ICI and stimulus-level configura-

tion (panel b). Hilbert envelopes were obtained by calculating the analytical signal for the DS and US waveforms, as demonstrated in Fig. 3.7. The Hilbert envelopes were

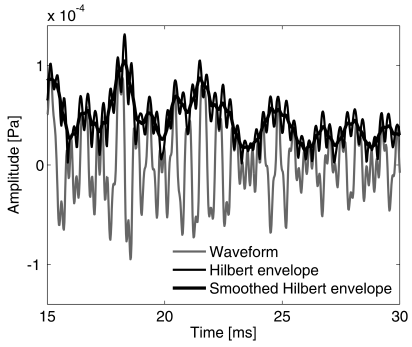


Figure 3.7: Envelope extraction method. Hilbert envelopes were extracted by calculating the analytical signal of the DS and US waveforms. The Hilbert envelopes were smoothed using a Savitzky-Golay FIR smoothing filter with order 3 and window-length of 51 samples to form the envelopes in Fig. 3.8 and Fig. 3.9. The technique was demonstrated on a 5-ms time window of the US waveform of subject #1 for L_t = 75 dB peSPL.

passed through a 3^{rd} order Savitzky-Golay FIR smoothing filter to form the envelopes in Fig. 3.8 and 3.9 (panel b).

The top panels of Fig. 3.8 and 3.9 show US waveforms of subjects #1 and #4 for increasing L_t. Subject #1 showed long-latency SSOAE components that rang until the end of the recording window, whereas the CEOAE of subject #4 disappeared in the noise floor of the recording after 35 ms. The almost tonal component, observed in the 26 to 36 ms window of the emission of subject #4, corresponded to the 1–1.5 kHz frequency band found in the long-latency spectrum in Fig. 3.4. It can be argued that this component represented an SSOAE. Both subjects showed overall increasing waveforms when L_t increased. Correspondingly, the rms levels of the short- and long-latency regions of these emissions grew compressively (see Fig. 3.5). There were no apparent waveform changes (other than magnitude) noticeable in the emissions of subjects #1 and #4 with increasing L_t. The zero-crossings of the CEOAEs remained roughly at the same locations for different intensities, though subject #1 showed a small latency decrease of the CEOAE to L_t = 60 dB peSPL compared to the CEOAEs to higher L_t. Carvalho et al. (2003) and Goodman et al. (2009) have investigated

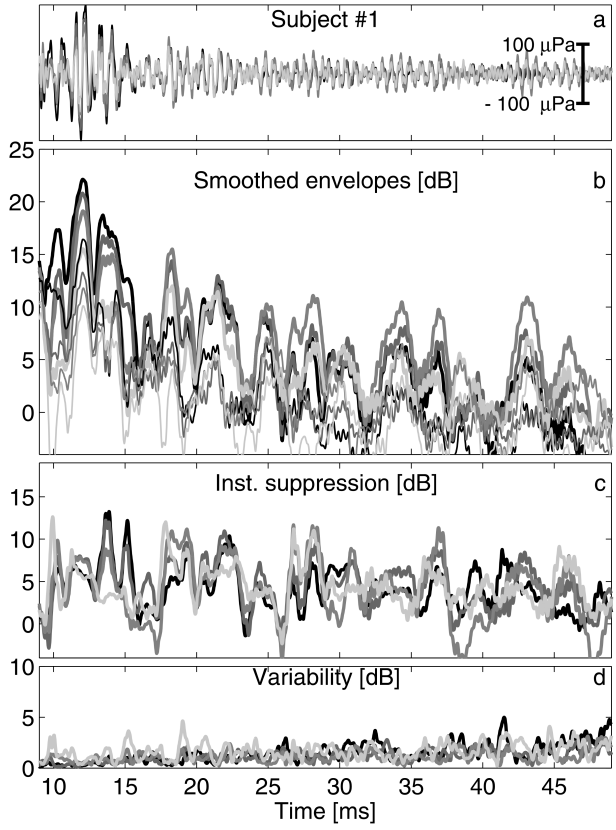


Figure 3.8: Envelopes and instantaneous suppression derived for subject #1 and L_t of 75 (black), 70, 65 and 60 (lightest gray) dB peSPL. Instantaneous suppression and DS envelopes were calculated for an ICI of 2 ms and $L_s = L_t$. a) Overlaid US waveforms for L_t = 75 to 65 dB peSPL in steps of 5 dB. b) Corresponding envelopes extracted with the method in Fig. 3.7 for US (thick) and DS (thin) waveforms. c) Instantaneous suppression calculated as the difference between corresponding US and DS envelopes of panel b. d) Variability of the instantaneous suppression measure.

latencies of several frequency components in the CEOAE to increasing intensities of the test click. Whereas Carvalho et al. (2003) found varying latencies for different band pass filtered frequency components across subjects, Goodman et al. (2009) did not find significant latency differences of frequency components in the "later" latency

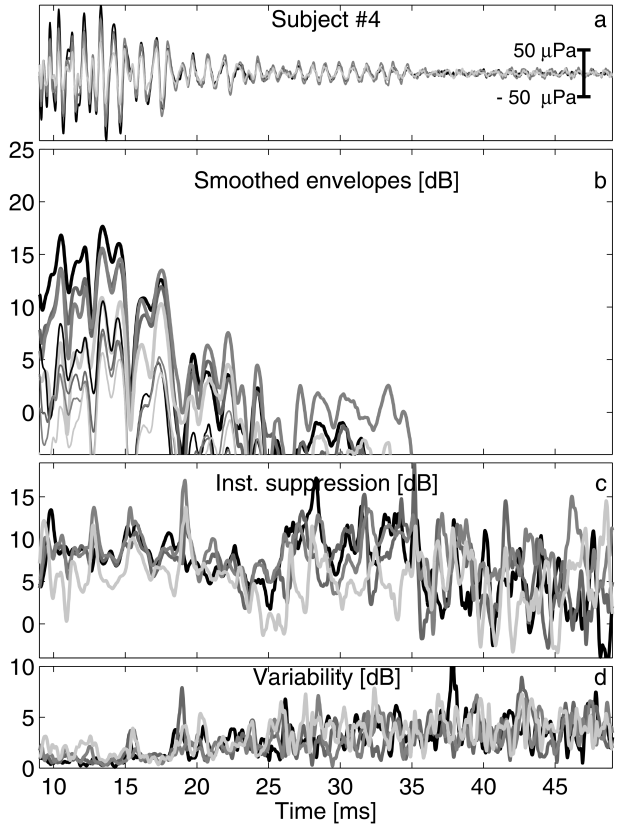


Figure 3.9: Envelopes and instantaneous suppression derived for subject #4 and L_t of 75 (black), 70, 65 and 60 (lightest gray) dB peSPL. Instantaneous suppression and DS envelopes were calculated for an ICI of 2 ms and L_s = L_t. a) Overlaid US waveforms for L_t = 75 to 65 dB peSPL in steps of 5 dB. b) Corresponding envelopes extracted with the method in Fig. 3.7 for US (thick) and DS (thin) waveforms. c) Instantaneous suppression calculated as the difference between corresponding US and DS envelopes of panel b. d) Variability of the instantaneous suppression measure.

(> 2.5 ms) components of the CEOAE with increasing intensity. The bandpass (1–3 kHz) filtered CEOAE envelopes of Carvalho et al. (2003) on the other hand, did not show phase shifts with increasing stimulus intensity. This was in agreement with the

US envelopes observed of Fig. 3.8 and Fig. 3.9, that did not show altered latencies for the peaks and troughs of the envelopes when L_t increased.

The DS envelopes plotted in panel b of Fig. 3.8 and 3.9 for an ICI of 2 ms and L_s = L_t showed similar peak and trough latencies than for the US envelopes. This shows that when a suppressor was presented, the envelopes of the waveforms were reduced in magnitude and not in waveform shape. This was one of the main requirements for the measurement paradigm in Fig. 3.2 to function well in describing changes in cochlear compression (Verhulst et al., 2008). To investigate whether there were differences in instantaneous suppression for the peaks and troughs of the waveforms, suppression was calculated as the difference between the DS and US envelopes of panel b. Instantaneous suppression fluctuated between -5 and 12 dB for subject #1 and between -2 and 17 dB for subject #4. The variability of the instantaneous suppression measure was calculated in panel d to verify how reliable these fluctuations were. The variability was calculated by deriving five DS and US envelopes that were obtained by averaging 360 DS and US responses in stead of averaging 1800 DS and US averages for every L_t in panel b. Instantaneous suppression was calculated for these 5 'sub' US and DS envelopes, and the standard deviation of suppression for these 'sub' envelopes thus represented the variability of the measure. For both subjects, the variability of the instantaneous suppression measure increased with increasing latency. The variability of subject #4 was systematically larger than 5 dB for latencies above 25 ms; subject #1 had a smaller variability in that same latency region. Both subjects showed a larger variability for the minima than for the maxima of the envelopes of US and DS in panel b. Taking into account the variability, instantaneous suppression patterns for both subjects showed showed similar features across L_t. Since Fig. 3.8 and 3.9 showed that the envelopes and instantaneous suppression patterns had similar shapes and latencies for different level configurations, it was justified to calculate suppression across fixed short- and long-latency windows of the waveforms in Fig. 3.3, 3.5 and 3.6.

For subject #1 in Fig. 3.8, the waveforms of instantaneous suppression followed the peaks and troughs of the US and DS envelopes in panel b. This was demonstrated for the suppression levels of the peaks and troughs for L_t = 75 dB peSPL. Subject #1 showed instantaneous suppression levels of 6.7, 5.6, 13, 2.3, 5.2, 7.9, and 4.7 dB for peak envelope latencies at 10.2, 12, 13.5, 16.6, 18.3, 21.5 and 24.4 ms, in the latency

region between 10 and 26 ms. The suppression levels for the troughs in that same latency region had levels of 4.6, 3.5, 0, 0.7, 2.2, -0.4 and -1.2 dB for latencies of 10.5, 12.6, 16, 17.5, 20.2, 23.6 and 26 ms. The average peak suppression was thus 6.5 dB, whereas the average through suppression was 1.3 dB. Even when a variability of 3 dB was taken into account for the suppression levels, it can be concluded that suppression for subject #1 was larger in the peaks of the DS and US envelopes than in the troughs. A similar peak and through analysis for instantaneous suppression of subject #4 in Fig. 3.9 did not demonstrate such a clear waveform dependency of instantaneous suppression. The variability near the troughs of the DS and US envelopes interfered with the instantaneous suppression measure, such that suppression for the peaks and troughs could not be compared for subject #4.

3.4 Discussion

This study has demonstrated that the CEOAE level-curve to a test click changed when a suppressor was presented 0–8 ms before the test click. This dynamic (i.e., time-dependent) change of CEOAE level-curve was designated temporal adaptation of the CEOAE level-curve. Temporal adaptation occurred for all four subjects, and for the short- and long-latency CEOAE. This indicated that temporal suppression affected SSOAE (i.e., long-latency region for subjects #1 and #2) and purely evoked CEOAE components (i.e., short-latency region for subjects #3 and #4). Subject-dependent differences were found in the shape of the unsuppressed CEOAE level-curve, and in the amount and time scale of temporal suppression. Temporal adaptation of the CEOAE level-curve was thus found to consist of a shared underlying component that was similar for all subjects, and a subject-dependent component that described the magnitude and time scale of suppression.

The results in this study challenge Kapadia and Lutman (2000b) who suggested that temporal suppression did not follow the compressive behavior of the CEOAE level-curve. Kapadia and Lutman (2000b) expected an increase of maximal suppression with increasing L_t if the level dependence of suppression were to follow the increase in compression of the CEOAE that is generally observed when increasing L_t (i.e. the unsuppressed level-curve). In stead, they found a non-monotonic increase of

maximal suppression with increasing L_t. Maximal suppression for an L_t of 50 and 60 dB ($\sim$6 dB) was found to be larger than for an L_t of 40 and 70 dB ($\sim$4 dB), when $L_s = L_t$. In this study, suppression was found to increase with increasing L_t, except for L_t from 70 to 75 dB peSPL, where the mean suppression decreased by 0.5 dB for an ICI of 2 and 3 ms. The suppressed level-curves in Fig. 3.5 and 3.6 furthermore showed linear growth below the knee-point, and compressive growth above, similar to the growth of the unsuppressed level-curves. There were two differences between the study of Kapadia and Lutman (2000b) and ours that might be the basis for the discrepancy found with respect to the CEOAE level-curve. First of all, Kapadia and Lutman (2000b) plotted the *maxima* of suppression patterns as a function of L_t, instead of plotting suppression for fixed ICIs, as in Fig. 3.5 and 3.6. The results elsewhere in Kapadia and Lutman (2000b) suggest that maximal suppression occurred for shorter ICIs when increasing L_t. Therefore, they might have observed increasing suppression with increasing L_t if they would have considered suppression for fixed ICI. Secondly, the mean suppression levels in Kapadia and Lutman (2000b) did not reveal information about the shape of the unsuppressed CEOAE level-curve, and about how it would alter when presenting a suppressor. In contrast, the suppressed level-curves in this study were analyzed on an individual basis, such that temporal adaptation of the CEOAE level-curve could be demonstrated.

3.4.1 Dynamic changes in local place-fixed BM mechanisms

The results from this study, together with the current knowledge on the generator mechanisms behind CEOAEs, have brought us closer to understanding the origin of temporal suppression. Withnell et al. (2008) and Goodman et al. (2009) have recently investigated the generator mechanisms behind the late-latency region of the CEOAE, which was set to start between 2.5 to 6.8 ms after click onset. The mechanisms underlying this late-latency region of the CEOAE were found to comply well with the coherent reflection theory (Zweig and Shera, 1995; Kalluri and Shera, 2007; Withnell et al., 2008; Goodman et al., 2009). The coherent reflection theory describes a scattering process in the cochlea that arises when a traveling wave encounters place-fixed BM impedance irregularities along its pathway [see Shera and Guinan (2007) for a comprehensive description]. Traveling wave scattering and coherent reflection are

essentially linear processes (Shera and Guinan, 1999), but the reverse traveling wave does reflect the compressive nonlinearity observed in the evoking forward traveling wave (Kalluri and Shera, 2007). Thus, the compression observed in the CEOAE level-curve is attributed to the characteristics of the forward traveling wave, and not to the actual reflection mechanisms that are the basis of reverse traveling waves in the cochlea (Kalluri and Shera, 2007). The results of this study attribute temporal adaptation of the CEOAE level-curve to changes in local place-fixed BM mechanisms, and there were two observations that supported this conclusion.

Firstly, temporal suppression was observed for the short- (6–18 ms) and long-latency (24–36 ms) CEOAE. These time windows were both situated in the late-latency region of Kalluri and Shera (2007), Withnell et al. (2008) and Goodman et al. (2009), and reflected either evoked emission components (short-latency) or SSOAE components (long-latency). Even though the short-latency region of subjects #1 and #2 consisted of both SSOAE and purely evoked components, the slope of the unsuppressed CEOAE level-curve decreased from the short- to long-latency region (Fig. 3.5). The OAE taxonomy of Shera and Guinan (1999) classified CEOAEs and SOAEs both as reflection source emissions, because they share the need for coherent reflection to occur. SOAEs were, in this taxonomy, explained as long-lasting amplitude stabilized cochlear standing waves that arise through narrow-band coherent reflection (Zweig and Shera, 1995; Shera and Guinan, 1999). The difference in compression, observed for the short- and long-latency region of the unsuppressed CEOAE-level curves of subjects #1 and #2, supported this classification. The saturation observed in the long-latency level-curve complied with the behavior of amplitude stabilized standing waves that would not increase in level after a certain stimulus threshold was reached. The evoked short-latency emission component on the other hand, reflected emission components that had not stabilized, and would thereby follow the compression characteristics of the evoking traveling wave better. The adapted level-curves for subjects #1 and #2 showed vertical offsets in the compressive region of the level-curves, and had similar compression slopes than the unsuppressed CEOAE level-curves (Fig. 3.5). Thus, even though the unsuppressed CEOAE level-curve changed considerably from the short- to long-latency region in Fig. 3.5, temporal suppression did not. The difference between amplitude stabilized cochlear standing waves and evoked emission

components were thus found to be well represented in the compression characteristics of CEOAE level-curve, but not in temporal suppression. This observation implies that temporal suppression is linked to local BM I/O functioning, and not to the mechanisms responsible for the specific features of these two types of emissions. If both the evoked and spontaneous emission types are similarly affected by temporal suppression, as the results in Fig. 3.5 and 3.6 suggest, then the origin of temporal suppression must lie in the mechanism that is common to both. Following the taxonomy of Shera and Guinan (1999), this would be the location where coherent reflection takes place, thus local place-fixed BM mechanisms.

A second observation in this study strengthens the statement that temporal suppression originated from local BM mechanisms. As shown in Fig. 3.8 and 3.9, instantaneous suppression was found to fluctuate across latency. For subject #4, the pattern of this fluctuation was biased by a large variability of instantaneous suppression in the dips of the waveforms. For subject #1 on the other hand, instantaneous suppression was clearly observed to follow the envelope of the CEOAE waveform for latencies between 10 and 26 ms. Shera and Cooper (2009) have investigated the mechanisms underlying specific features of the CEOAE waveform. They found that the waxing and waning of the CEOAE (i.e., the multiple temporal lobes that are often observed in the CEOAE waveform) were related to multiple reflections patterns of CEOAE energy inside the cochlea. They related the waxing periods of the CEOAE to place-fixed BM irregularities, and the waning periods to internal reflections of the reverse traveling wave to the stapes. This reflected energy would in turn excite the BM again, whereby a second waxing period in the CEOAE waveform would be created. Instantaneous suppression for subject #1 was related to the waxing and waning of the OAE by showing greater suppression levels for the peaks of the CEOAE envelopes than for the troughs. The correspondence between increased suppression for waxing periods, and decreased suppression for waning periods, suggests that suppression is related to local BM mechanisms. If the waxing periods of the CEOAE are related to place-fixed BM irregularities, as suggested by Shera and Cooper (2009), then this would explain the waveform dependency observed in instantaneous suppression.

3.4.2 Relation to BM impulse response data

There is no physiological data available that investigates BM mechanics to the double-click stimuli that were used in this study, but BM responses to single clicks were measured by Recio and Rhode (2000) for chinchilla. Even though BM click-response data cannot give more information about adaptive processes that take place inside the cochlea when presenting two clicks close in time to each other, they do contain valuable information about frequency and level dependencies of the BM impulse response (IR). Recio and Rhode (2000) showed that BM IRs in chinchilla were shorter ($\sim$1.2 ms) at high-frequency locations (13.7 kHz) than at low-frequency locations ($\sim$2.2 ms at 6.1 kHz). They furthermore showed that BM IR maxima occurred later for low-frequency BM locations than for high-frequency BM locations ($\sim$ 0.5 ms later for 6.1 than for 13.7 kHz). It was investigated here whether the frequency and level dependencies of the BM IRs found by Recio and Rhode (2000) were in any way related to frequency and level dependencies in temporal suppression.

Making a link between features of BM IRs and temporal suppression is only valid if temporal suppression is related to the duration of local BM IRs. It is postulated here that, regardless of whether a static or dynamic nonlinearity underlies temporal suppression, the effect would be related to the duration of the local BM IR. In the case of a static nonlinearity, suppression would only occur as long as two BM click responses overlap in time. This theory is based on the idea that temporal overlap of BM impulse responses would change the input to a static nonlinearity (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Hine and Thornton, 2002; Harte et al., 2005), and thereby change the amount of local BM compression. Constructive summing of BM impulse response patterns to two closely spaced clicks would lead to larger inputs to the static nonlinearity than for a single click, and this would be observed as suppression when applying the measurement paradigm of Fig. 3.2 (Kapadia and Lutman, 2000a; Harte et al., 2005). The idea behind a dynamic nonlinearity underlying temporal suppression, is that the applied compression for the second click would depend on the BM impulse response of the first click (Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Verhulst et al., 2008). This dynamic change in compression would be related to adaptation processes in the local nonlinearities, perhaps outer hair cell adaptation (i.e. the K^+/Ca^{2+}

transduction channel in the stereocilia of the hair cell can be saturated when deflecting the hair cell from its resting state (Hudspeth and Gillespie, 1994)). This adaptation process is expected to take place during the time course of the BM impulse response of the first click.

For both nonlinearity strategies (i.e., static or dynamic), temporal suppression is thus expected to be related to the BM impulse response duration. Hence, there might exist a relation between the frequency and level dependencies in the BM impulse response data of Recio and Rhode (2000), and the frequency components in the individual CEOAEs that contributed to temporal suppression. Based on the BM impulse response duration of Recio and Rhode (2000), it is expected that the time scale of temporal suppression for an exemplary case where the CEOAE would consist of low-frequency components (e.g. 1–2 kHz) would be longer than for a CEOAE with a high-frequency (e.g. 3–4 kHz) content only. Because human CEOAEs are generally broadband, it is difficult to demonstrate such a frequency dependency in temporal suppression; however, the temporal suppression patterns of the short- and long-latency regions of subject #4, may point to a such a relation (Fig. 3.3). The spectrum of subject #4 was observed to change from broadband (1–4 kHz) in the short-latency region to narrowband (1–1.5 kHz) in the long-latency region (see Fig. 3.4). At the same time, the width of the temporal suppression patterns in Fig. 3.3 for this subject became broader, and maximal suppression was found for higher ICIs. Whereas suppression for the short-latency pattern of subject #4 would be a combination of several locations along the BM, and would reflect high (shorter) and low (longer) frequency BM IRs, suppression for the long-latency CEOAE of this subject would only reflect contributions from the 1–1.5 kHz region. The change to a more broad suppression pattern for the long-latency CEOAE might be due to the loss of the shorter, high-frequency BM IRs that contributed to temporal suppression. Secondly, the shift of maximal suppression to higher ICIs for subject #4, when going from the short- to long-latency CEOAE, might also have been due to the loss of short, high frequency BM IRs that contributed to temporal suppression. The maxima of the BM IR data of Recio and Rhode (2000) were found to occur later for the low than for the high frequencies. Thus, the loss of high-frequency content in the long-latency CEOAE of subject #4 may have caused maximal suppression to occur at longer ICIs.

An observation from the data of Recio and Rhode (2000) revealed that BM IRs kept their temporal fine-structure when increasing stimulus intensity (Shera, 2001a). However, the maxima of the BM impulse responses occurred earlier when increasing stimulus intensity. The envelopes of BM IRs thus changed with intensity, while their temporal fine-structure was invariant. This would imply that reductions in L_s for a fixed ICI and L_t, would be observed as a magnitude effect in the local BM IR, and not as a fine-structure change. Also, the BM IR maxima of the suppressor click would be delayed when decreasing L_s. At the same time, the BM IR to the test click would remain at the same location. The difference in maximal BM IR overlap between the patterns of L_t and L_s could cause the overlap (static nonlinearity) or adaptation effect (dynamic nonlinearity) for the BM impulse response envelopes to be maximal at a later time for lower L_s. If this effect would occur for all ICIs and all BM locations that contribute to the CEOAE, then a shift of maximal temporal suppression to higher ICIs would be expected when reducing the L_s relative to L_t. The data in Fig. 3.3 does not show a pronounced shift of maximal suppression to longer ICIs with decreasing L_s, but the short-latency data of subjects #1, #2 and #4 does point in that direction. Comparably, the suppression maxima in Kapadia and Lutman (2000b) showed a clear shift from 2.8 ms to 0.8 ms when increasing L_t from 40 to 70 dB, while keeping L_s = L_t. There thus seems to be a relation between the delayed suppression maxima and the delayed BM impulse response maxima observed for lower click intensities (Recio and Rhode, 2000). Even though the attempt to link level and frequency dependencies of temporal suppression to BM IRs was only speculative, our results did not contradict BM impulse response behavior observed in animal data.

3.4.3 Static versus dynamic nonlinearity

Neither the place-fixed cochlear origin of temporal suppression, nor its frequency and level dependence deducted from BM IRs, revealed information about the nature of the local BM nonlinearity that underlies temporal suppression. Kapadia and Lutman (2000b) and Harte et al. (2005) modeled compression changes for one location on the BM (i.e., single channel processing) by presenting temporally spaced 1 kHz 10-ms tonebursts to a static nonlinearity (i.e., automatic gain control system with a compressive nonlinearity) that represented a local I/O function. The input to this local

BM nonlinearity would change depending on the temporal overlap of the tonebursts. When applying the same experimental paradigm as in Fig. 3.2, this led to a model that accounted for suppression as long as the tonebursts overlapped in time. Even though the models of Kapadia and Lutman (2000a) and Harte et al. (2005) only considered single channel processing, they accounted well for reduced suppression for increasing ICIs. Conversely, temporal suppression maxima for ICIs different than 0 were not adequately accounted for. One improvement to their model would be to use stimuli that are different than their symmetric tone bursts. For example, asymmetric tonebursts that contain similar features as the BM IRs observed in Recio and Rhode (2000), could be used. The tonebursts could be designed to keep a temporal fine-structure invariance, while changing their maximum envelope to shorter durations for increasing intensities. Implementing asymmetries in the input signals to the static nonlinearity might be able to account better for the asymmetries observed in the temporal suppression pattern. Nevertheless, even if these model improvements are made, a single channel model would not be able to account for across channel or broadband cochlear effects. Future research of these local single channel models, all with different BM IR features, might be able to account for temporal adaptation of the CEOAE level-curve.

Alternatively, Verhulst et al. (2008) presented a phenomenological model for temporal adaptation of the CEOAE level-curve. They suggested that the suppressed CEOAE level-curve would show a dynamic shift in knee-point position depending on ICI, such that suppression and augmentation (i.e., negative suppression) would both be accounted for. The knee-point of the suppressed CEOAE level-curve was suggested to shift to lower stimulus intensities to account for suppression, and to higher stimulus levels to account for augmentation. The underlying mechanisms for suppression and augmentation were argued to stem from dynamics in local BM nonlinearities, that would result in a suppressed level-curve that varied depending on the ICI (Verhulst et al., 2008). Lina-Granade and Collet (1995) and Kapadia and Lutman (2000b) brought up the possibility of local dynamic changes in BM I/O functions before, but did not find a strong argumentation for them to be necessary. Verhulst et al. (2008) argued that static BM I/O functions alone would not be able to account for the augmentation they observed in their experimental data. The observation of temporal

augmentation has been debated in the past, and the study of Verhulst et al. (2008) was the first to demonstrate augmentation for the measurement paradigm in Fig. 3.2. Conversely, there were no observations of augmentation in the results presented in Fig. 3.3, 3.5 and 3.6. The ambiguity between the results in Verhulst et al. (2008) and this study was investigated in Table A.2 of the Appendix. It was shown that augmentation can arise as an artifact of the measurement implementation of Verhulst et al. (2008). The interleaved recording technique adopted in this study eliminated this measurement bias, and thereby showed more reliable suppression results than Verhulst et al. (2008). Verhulst et al. (2008) favored a dynamic over a static nonlinearity based on the observation of temporal augmentation. Even though their argumentation seems not to hold on the observation of augmentation alone, the concept of a dynamic nonlinearity remains applicable for the results found in this study. The phenomenological model of Verhulst et al. (2008) predicted the results in Fig. 3.5 and 3.6 better for the compressive than for the linear region of the CEOAE level-curve. Whereas the predicted knee-point location of the suppressed CEOAE level-curve in Verhulst et al. (2008) changed with ICI, no pronounced shifts were observed in the knee-point location in Fig. 3.5 and 3.6.

As temporal adaptation of the CEOAE level-curve showed features that can be related to both static and dynamic nonlinearities, the data are inconclusive as to which of the two local BM mechanisms (or a combination of both) might underly temporal suppression. To find the origin of temporal suppression, a link between the dynamics of local BM functioning, and dynamic changes of the broadband CEOAE level-curve should be established. One way to investigate the BM mechanisms underlying temporal suppression would be to record animal BM impulse responses to closely-spaced clicks. Alternatively, it would be useful to simulate the obtained results with a model of the cochlea that can account for across channel (i.e., broadband) processing. Time-domain models of the cochlea, such as transmission-line models, would be well suited to simulate BM impulse responses and excitation patterns to the stimuli used in this study.

3.5 Conclusion

This study demonstrated temporal adaptation of the CEOAE level-curve when a suppressor click was presented 0–8 ms before a test click. Temporal suppression patterns were obtained for ICIs between 0 and 8 ms. Temporal suppression occurred for all subjects, but subject-dependent differences in the magnitude of suppression, the width of the suppression pattern, and the location of suppression maxima were found. Temporal suppression and adaptation were observed for the short- and long-latency region of the CEOAE, and affected the CEOAE latency regions dominated by evoked components and SSOAE components similarly. The origin of temporal suppression was demonstrated to be linked to place-fixed BM I/O functions, but the results were inconclusive about the type of BM nonlinearity (static or dynamic) that would best describe the results. A more comprehensive relation between temporal characteristics of local BM I/O functions and the dynamics of the broadband CEOAE level-curve should be established to better understand the cochlear origin of temporal adaptation of the CEOAE level-curve.

4

Modeling Otoacoustic Emissions with a Transmission-Line Model of the Human Cochlea

To investigate the underlying cochlear mechanisms of temporal adaptation in click-evoked otoacoustic emissions, a model approach is needed. The frequency components in the CEOAE stem from several locations along the basilar membrane, hence the CEOAE describes a broadband cochlear response. To model CEOAEs, it is desirable to use a model that describes the whole length of the cochlea. One such candidate is a transmission-line model of the cochlea that describes traveling waves along the length of the BM, and that can be adapted to produce otoacoustic emissions. This chapter describes the development of a transmission-line model of the cochlea that accounts for forward and backward traveling waves (i.e., OAEs). The model was modified such that it was capable of simulating both reflection source (i.e., CEOAEs and stimulus-frequency OAEs) and distortion source emissions (i.e., distortion-product OAEs). Reflection source emissions are accounted for by including irregularities in the impedance function describing the BM mechanics. The generation of distortion source emissions is made possible by implementing a nonlinearity in the local BM impedance functions that describe the gain properties in local BM input/output functions. This chapter is concluded by evaluating the performance of the model in simulating key properties of cochlear functioning such as cochlear dispersion, level-dependent features of local BM impulse responses, and the sharpness of auditory filters.

4.1 Introduction

Transmission-line models have been widely adopted to investigate cochlear function, and became popular in the beginning of the eighties (e.g. de Boer, 1980; Kim et al., 1980b). Transmission-line models represent the cochlea as a cascade of a large number of band-pass filters that are tuned to different characteristic frequencies, corresponding to the tonotopic organization of the BM (i.e., high frequencies at the base and low frequencies at the apex; Greenwood, 1961). The advantage of transmission-line models over "functional" cochlear models such as the gammatone filter bank (de Boer, 1975) and the dual-resonance nonlinear (DNRL) filter bank (Lopez-Poveda and Meddis, 2001) is that they simulate biophysical properties of the BM. Rather than describing a bank of independent filters fitted to the tuning characteristics of auditory nerve fibers, transmission-line models account for the local mechanics of the BM by implementing a spatial impedance function. Not only does this approach lead to a set of coupled filters, it makes it possible to simulate traveling waves. The success of transmission-line models of the cochlea is partly due to the ease with which the impedance function can be adapted to include either passive or active (i.e., internal energy sources) elements. There is a consensus regarding OHCs being able to pump energy into BM vibration, which advocates for active elements in the transmission-line. However, it is unclear how this coupling of OHC amplification to the BM is best represented, and transmission-line models provide an excellent platform to investigate this. Transmission-line models can furthermore be made to account for backward traveling waves or otoacoustic emissions (e.g. van Hengel, 1996), which makes them interesting for this project. Transmission-line models can be solved over time so that the BM velocity along the whole length of the cochlea is known for every time step of the stimulus (Diependaal et al., 1987). This makes them suitable for investigating time-dependent features of cochlear processing.

One-dimensional transmission-line models represent the uncoiled cochlea typically as a closed rectangular box. The BM is positioned centrally along the length of the box and surrounded by incompressible fluid that represents scala vestibuli and scala tympani. This simplified representation of the cochlea was justified by the early measurements of von Békésy (1960) on human cadaver cochleae. It is commonly

assumed that the wavelengths of the BM waves are large compared to the height of the scalae in each of the cochlear cross sections. This assumption is also known as the "long-wave approximation", and ensures that the pressure is uniformly distributed in the perpendicular directions of the BM.

Two important examples of linear transmission-line models were proposed by de Boer (1980) and Zweig (1991). Both models represented the BM as a cascade of coupled mass-spring-damper systems, and described traveling waves as a consequence of the fluid coupling between the different resonating sections of the BM. A fundamental difference between the two approaches is found in how they represent the cochlear input impedance Z_0 (i.e., the load impedance seen by the middle-ear). Both de Boer (1980) and Zweig (1991) based their spatial BM impedance function on BM vibrations measured at different locations along the BM (e.g. Rhode, 1971). To evaluate their implementations in terms of the coupling between the cochlear fluids (i.e., cochlear impedance) and the BM (i.e., cochlear admittance), the input impedance of their models should agree with measurements of Z_0 of the cochlea. The implementation of de Boer (1980) corresponded to a capacitive Z_0 (i.e., negative phase angle $\angle Z_0$ increased to zero with frequency), whereas the implementation of Zweig (1991) led to a resistive Z_0 (i.e, constant $\angle Z_0$ across frequency). The measurements of Nedzelnitsky (1980) and Lynch et al. (1982) support the view of a resistive over a capacitive Z_0, which meant that an adaptation to the model implementation of de Boer (1980) was needed to account for this. In the model of de Boer (1980), the cochlear admittance varied along the length of the BM by increasing the compliance across the BM, while the acoustic mass of the cochlear fluids (i.e., cochlear impedance) across the BM was assumed constant. The combination of the cochlear impedance and admittance thus led to a capacitive Z_0 in the model of de Boer (1980). Shera and Zweig (1991) proposed that a proportionality between the spatial cochlear impedance and admittance could account for a resistive cochlear Z_0. This proportionality stemmed from the reduced dimensions of the cochlea when going from base to apex (i.e., tapering). Zweig (1991) implemented this relation between the cochlear impedance and admittance by maintaining a proportionality between the acoustic mass of the cochlear fluids and the capacitance of the BM (Shera and Zweig, 1991). The advantage of a resistive over a capacitive Z_0 is that the wavelength for the resistive Z_0 changes slowly near the base

of the cochlea, such that traveling waves undergo little reflections in either direction (Shera and Zweig, 1991). Breaking the cochlear symmetry (i.e., the spatially varying cochlear impedance and admittance) can furthermore lead to erroneous low-frequency standing waves in Z_0 (Puria and Allen, 1991; Shera and Zweig, 1991).

To account for a tall and sharp peak in BM transfer functions (i.e., auditory filters) that show a broad base, active elements need to be implemented in the transmission-line model. Passive models can create relatively sharp transfer functions by decreasing the damping in the model, but do not account for the transfer-function peak height and base width observed in experimental data (Zweig, 1991). A commonly adopted technique to make the model active, is to include negative damping in the spatial BM impedance function such that energy is fed into the traveling wave (e.g., Zweig, 1991). The idea of using negative damping stems from the functioning of the OHC, that are able to feed energy to the BM vibration pattern through OHC motility. OHC motility (i.e., the change in OHC length when the voltage across its membrane is changed Brownell et al., 1985), has been implemented in transmission-line models either as OHC feedback forces (e.g., Geisler, 1991; Neely, 1993; Liu and Neely, 2009; Moleti et al., 2009), or as limit-cycle oscillators (e.g., Duifhuis et al., 1986; Talmadge et al., 1998).

The model presented here was based on the linear transmission-line model of Zweig (1991). This model accounted for sharp auditory filters by including negative damping, and furthermore implemented tapering of the cochlea by making both the cochlear impedance and admittance dependent on cochlear location (Shera and Zweig, 1991). The model was derived on the basis of a local scaling symmetry that has been observed in measured BM transfer functions (Rhode, 1971; Rhode and Cooper, 1996). Local scaling symmetry implies that cochlear transfer functions scale with $f/f_{cf}(x)$, a parameter that connects position x and frequency f through the cochlear map $f_{cf}(x)$ (Zweig, 1976). By using this principle, Zweig (1991) was able to describe vibration patterns along the length of the cochlea, by describing the mechanics at one location along the BM as a function of the scaling variable $f/f_{cf}(x)$. The transmission-line equations by Zweig (1991) were in this study applied to a large number of BM sections such that a set of model equations, each describing the mechanics at a specific f_{cf}, were obtained. Solving this set of equations allowed for the simulation of a traveling

wave along the length of the BM. Because the purpose of this study was to investigate dynamic features of cochlear compression by using otoacoustic emissions, the model was adjusted in two ways. Firstly, the model was solved over time, such that time-dependent properties of the traveling wave and cochlear compression could be investigated. The fourth order Runge-Kutta method, that was earlier adopted by Diependaal et al. (1987) to solve transmission-line models over time, was implemented for this purpose. Secondly, reverse traveling waves or otoacoustic emissions were generated by including random fluctuations on top of the negative damping function along the length of the cochlea. Zweig and Shera (1995) have shown that coherent reflections can take place at the sites of cochlear irregularities when a traveling wave encounters these irregularities. As a consequence, backward traveling waves or OAEs are created. Reflection source emissions such as CEOAEs and SFOAEs were simulated by implementing these irregularities, however this did not account for the nonlinear characteristics of these emissions. The compression observed in the level-curve of CEOAEs is associated with nonlinear properties of the OHC (Kemp, 2007), and thus the linear model implementation was extended to include nonlinear properties of the BM input/output function (e.g., Robles et al., 1986).

Including BM nonlinearities in the model made it possible to simulate CEOAEs that showed increased compression for increased stimulus intensities, and made it furthermore possible to simulate distortion source emissions such as DPOAEs. Unlike several other time-domain models of the cochlea (van Hengel, 1996; Talmadge et al., 1998; Elliott et al., 2007; Moleti et al., 2009), the developed model takes a functional approach rather than a micro-mechanical approach for the nonlinearity. This choice was based on the application of the model for investigating the nonlinear mechanisms behind CEOAEs. Rather than making assumptions of how the micro-mechanics of OHC motility change along the length of the BM, a functional approach that accounted for the nonlinearities observed in BM impulse responses was used. The magnitude and envelopes of BM impulse responses (at one BM location) have been shown to change as a function of stimulus-intensity, whereas their temporal fine-structure was shown to be essentially unaltered (Recio and Rhode, 2000). These results were accounted for in the model of Shera (2001b) by moving the poles of the BM impedance of Zweig (1991) vertically in the phase-pole diagram, depending on the stimulus inten-

sity. Shera's model was able to account for the increased compression and invariance of the zero-crossings observed in the BM impulse responses recorded to increasing stimulus intensities (Recio and Rhode, 2000). The nonlinearity strategy was furthermore shown to account for frequency glides observed in the experimental data (Shera, 2001a). The functional model of Shera (2001b) was adapted to describe all sections of the BM at once, and was implemented in the time-domain model developed here.

The nonlinear transmission-line model of the cochlea that is presented in this study was based on existing concepts developed by Diependaal et al. (1987), Zweig (1991), Shera and Zweig (1991) and Shera (2001b). Therefore, this study did not contribute new nonlinearity strategies or model ideas; it combined and validated well established work. The purpose of the model was to account for nonlinearities observed in BM impulse responses, while at the same time satisfying level-dependent features of the CEOAE. If the model is viable, then it can be adopted to investigate dynamic features of cochlear nonlinearities underlying the dynamic changes observed in the CEOAE level-curve.

4.2 Description of the linear and tapered transmission-line model

This section explains how the transmission-line equations of Zweig (1991) describing the mechanics of a section of the BM in function of the scaling variable $f/f_{cf}(x)$, were implemented in a transmission-line model that simulates traveling waves. Firstly, the equations of Zweig (1991) were written as a function of $f_{cf}(x)$ such that an equation was obtained for every section of the cochlea. Secondly, two boundary equations were developed: one for the connection of the cochlea to the middle ear, and one for the apex boundary. After a description was found for all sections of the cochlea, tapering was included by making the cochlear impedance proportional to the cochlear admittance. Lastly, the model was adjusted to simulate reflection source emissions. The format of the transmission-line equations was chosen to agree with earlier implementations of van Netten and Duifhuis (1983), Duifhuis et al. (1986) and van Hengel (1996), such that a framework that incorporates the different cochlear

models is possible. An acoustical representation of the transmission-line model is shown in Fig. 4.1. The volume velocity U_n at node n was given by $A\,\dot{y}$, and p_n described the force per unit area F/A. To obtain equations for every section of the

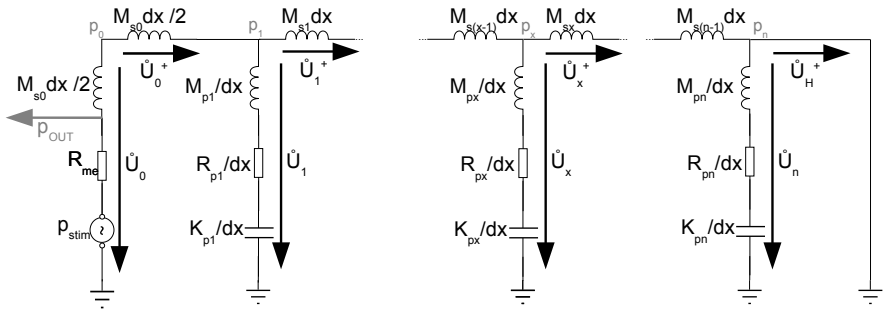


Figure 4.1: Schematic of the acoustical representation of the transmission-line model. U_n and p_n describe the volume-velocity and pressure for sections n of the model. M, R and K represent acoustical impedances of a mass, damper and spring respectively.

model, Newton's second law was applied. For an acoustical system, Newtons' second law is given by:

$$p = M\dot{U} + RU + KY\,, \tag{4.1}$$

where M, R and K represent the acoustical impedance of a mass, damper and spring respectively. For a mechanical system, this becomes:

$$p = m\ddot{y} + r\dot{y} + ky, \tag{4.2}$$

where y represents displacement, and where m, r and k represent mechanical elements (i.e. acoustical elements times unit area).

4.2.1 Solving the traveling wave equation

A description of the traveling wave along the length of the cochlea was found by solving the transmission-line schematic in Fig. 4.1 for every section n. The transmission-

line description of Zweig (1991) was used for this purpose. The following expressions for the pressure difference p over one section of the cochlea dx in function of the volume velocity U were given (Zweig, 1991):

$$\frac{dp}{dx} = -Z_s U \,, \tag{4.3}$$

$$\frac{dU}{dx} = \frac{-p}{Z_p} \,. \tag{4.4}$$

In eq. 4.3 and 4.4, Z_s represents the series or longitudinal impedance and Z_p the parallel shunt impedance of the cochlea. Zweig (1991) derived specific equations for the series and shunt-impedance based on BM transfer functions (i.e., BM velocity/umbo velocity) measured by Rhode (1971). In his model, Zweig (1991) assumed that the cochlea was represented by an uncoiled fluid-filled tube that contained the BM surrounded by incompressible fluids. The BM was represented by an array of beams that were coupled together through the fluids in the scalae. Zweig (1991) furthermore assumed that the pressure was uniformly distributed in the perpendicular directions to the BM, by applying the long-wave approximation (i.e., assumption that the wavelength of the traveling wave is large compared to the height of the cross-sections of the scalae at any location along the BM). Zweig (1991) furthermore used the approximate local "scaling symmetry" that is present in the cochlea to reduce the number of independent variables in his transmission-line description. The data of Rhode (1971) suggest that BM envelope responses to pure tones measured at one location can be translated into the envelope of the response at another cochlear location through a scaling of the envelope height. Using this scaling symmetry allows for a description of the cochlea in the scaling parameter $s = f/f_{cf}(x)$. As a consequence, the BM transfer functions along the length of the cochlea can be described by describing one transfer function measured at CF. Zweig (1991) described the BM transfer functions measured by Rhode (1971) and found the following expressions for Z_s and Z_p in Eq. 4.3 and 4.4:

$$Z_s(s) = \omega_{c0}\, M_{s0}\, s \,, \tag{4.5}$$

$$Z_p(s) = \omega_{c0}\, M_{p0} \left(s^2 + \delta s + 1 + \rho \mathrm{e}^{-\psi s}\right)/s \,. \tag{4.6}$$

The series impedance Z_s was dominated by the mass of the scalae at the base, M_{s0}. ω_{c0} represented the angular characteristic frequency at the base of the cochlea and reflected the highest frequency that was processed in the cochlea. The scaling parameter s was defined as:

$$s = \frac{j\omega}{\omega_{cn}}\,, \tag{4.7}$$

and depended on the angular characteristic frequency ω_{cn}. Z_p contained a damping constant δ and a stiffness term $1 + \rho e^{-\psi s}$. The stiffness contained a feedback term that depended on the BM displacement at specific time $\psi = 2\pi\mu$ before the current displacement. The feedback term was necessary to stabilize the unstable oscillations that would occur when choosing a negative δ. Near CF, a negative δ was necessary to account for the sharp peaks observed in the BM transfer functions by Rhode (1971). From the dimensions of Z_p in Eq. 4.6, it follows that the BM mass at the base, M_{p0}, has dimensions:

$$[M_{p0}] = \left[\frac{M}{L^3}\right]\,. \tag{4.8}$$

Equivalently, the dimensions for M_{s0} were given by:

$$[M_{s0}] = \left[\frac{M}{L^5}\right]\,. \tag{4.9}$$

The dimensions of the series and parallel mass were not identical, but as can be observed in the schematic in Fig. 4.1 and equations Eq. 4.3 and 4.4, this was accounted for by the multiplication with dx in Z_s, and the division by dx in Z_p. The equations for $Z_s(s)$ and $Z_p(s)$ developed by Zweig (1991), were rewritten such that an expression was found for every section (2000) of the cochlea. By doing this, it was possible to describe a traveling wave along the length of the modeled cochlea. Firstly, the shunt impedance Z_p was rewritten into an expression for the pressure p_n acting on a specific section of the BM in the schematic in Fig. 4.1. To couple the series to the parallel impedance, a pressure-balance along the length of the cochlea was written by applying Newton's second law in the horizontal direction. The pressure p_n depended on the pressure p_{n-1} acting on the previous section, and determined the pressure of the following section p_{n+1}. The traveling wave equation was then found by expressing the pressure differences across the different model sections in function of the BM velocity

and displacement. The specifics of this derivation can be found in Appendix B, and led to a transmission-line equation for every section n of the cochlea:

$$-\,q_{n-1} + A_{sc}\,q_n - q_{n+1} = A_{sq}[D\,\dot{y}_n + S\,y_n]\,, \tag{4.10}$$

where q_n described the pressure p_n per unit mass of the BM:

$$q_n = \frac{\omega_{cn}}{b\,\omega_{c0} M_{p0}}\,p_n \text{ with } [q_n] = \left[\frac{L}{T^2}\right]\,. \tag{4.11}$$

The damping D and stiffness S in Eq. 4.10 were given by:

$$D = \omega_{cn}\delta \text{ with } [D] = \left[\frac{1}{T}\right] \text{ and} \tag{4.12}$$

$$S = \omega_{cn}^2\left(1 + \rho \mathrm{e}^{\left(\frac{-j\omega\psi}{\omega_{cn}}\right)}\right) \text{ with } [S] = \left[\frac{1}{T^2}\right]\,. \tag{4.13}$$

Note that both D and S varied along the length of the cochlea and both decreased towards the apex of the cochlea if the damping constant $\delta > 1$ (i.e., passive model implementation). If $\delta < 1$, D increased exponentially towards zero from the base to the apex (i.e., active model implementation). The dimensionless variables A_{sq} and A_{sc} in Eq. 4.10 were given by:

$$A_{sq} = \frac{\omega_{cn}\,M_{sn}\,dx^2}{\omega_{c0}\,M_{p0}}\,, \text{ and} \tag{4.14}$$

$$A_{sc} = 2 + A_{sq}\,, \text{ with} \tag{4.15}$$

$$[A_{sq}] = [A_{sc}] = \left[\frac{1}{T}\frac{M}{L^5}L^2T\frac{L^3}{M}\right] = [\,]\,. \tag{4.16}$$

A_{sq} and A_{sc} described the ratio between the mass of the cochlear fluids times the angular critical frequency, to the BM mass times the angular frequency at the base of the cochlea. Both A_{sq} and A_{sc} varied along the length of the cochlea. An expression for the cochlear traveling wave was found by solving a matrix that consisted of n-2 transmission-line equations of the form of Eq. 4.10. The first two ($n = 0$ and $n = 1$) and last ($n = n$) rows of this transmission-line matrix were different

from the equation given in Eq. 4.10, because they represented the boundary of the cochlea at the helicotrema and at the middle ear. An expression for the middle-ear and helicotrema boundary was found by applying Newton's second law to the first and last section of the schematic in Fig. 4.1, and specifics on these derivations can be found in Appendix B.

4.2.2 The middle-ear boundary

The middle-ear boundary of the model describes the connection of the cochlea with the outside world and influences the traveling wave twofold. Firstly, the middle ear translates the stimulus pressure to a pressure applied to the base of the cochlea. The middle-ear transduction can be chosen to represent the human ear-canal and middle-ear impedance, such that the pressure applied to the base of the cochlea is bandpass filtered in the 2–4 kHz region. Alternatively, the middle-ear boundary can be chosen to represent a resistance that represents the cochlear input impedance near the base. In this case, the pressure at the base of the cochlea represents the stimulus pressure applied to the stapes (i.e., last middle-ear bone). Secondly, the choice of middle-ear boundary influences the reverse traveling wave generated in the cochlea. Any impedance mismatch between the middle ear and cochlear impedance at the base of the cochlea introduces reflections of the reverse traveling wave. These reflections can give rise to SOAEs that occur in the cochlea as standing waves caused by multiple internal reflections of energy (Shera, 2003). The amount of reflectivity present at the middle-ear boundary of the cochlea thus influences the generation and strength of the SOAEs in the model. The middle-ear boundary implemented in this model did not introduce a large reflectivity, and thus did not account for strong SOAEs. Rather than implementing a human middle-ear and ear-canal impedance, it was decided to apply the stimulus pressure to the base of the cochlea. These choices led to the following shunt-impedance Z_{p0} or Z_{me} at the base of the cochlea:

$$Z_{p0} = Z_{me} = \frac{-p_0}{U_0}\,. \tag{4.17}$$

The middle ear consisted of an acoustical mass M_{me} which equaled $M_{s0}\,dx/2$, an acoustical resistance R_{me}, and a pressure source p_{stim}. R_{me} was chosen to represent the cochlear input impedance at low frequencies. This "matching impedance" resulted in little reflections of energy of the backward traveling waves when they reached the base of the cochlea. R_{me} was described by:

$$R_{me} = \sqrt{\frac{Z_s}{Z_p}} \text{ at n=0}\,, \tag{4.18}$$

such that

$$R_{me} = \sqrt{M_{s0}\,K_0}. \tag{4.19}$$

The spring constant K_0 was given by Zweig (1991) as:

$$K_0 = \omega_{c0}^2\,M_{p0}\,, \tag{4.20}$$

such that R_{me} became:

$$R_{me} = \sqrt{\omega_{c0}^2\,M_{p0}\,M_{s0}} \text{ with } [R_{me}] = \left[\frac{M}{TL^4}\right]\,. \tag{4.21}$$

To integrate the middle-ear boundary into the transmission-line equations that were derived for the other sections of the cochlea in section 4.2.1, an expression was derived for the pressure balance at the middle-ear boundary. This expression was found by applying Newton's second law to the first section of the schematic in Fig. 4.1:

$$2q_0 - q_1 = \frac{p_{stim}}{M_{p0}\,b} + \frac{R_{me}\,dx}{M_{p0}}\,\dot{y}\,. \tag{4.22}$$

While the pressure balance at all sections of the cochlea is described by Eq. 4.10, the pressure balance at section 1 is represented differently due to the acoustical series mass $M_{s0}\,dx/2$, that was placed between p_0 and p_1 (see Fig. 4.1):

$$-\,2q_0 + A_{sc1}\,q_1 - q_2 = A_{sq1}[D\,\dot{y}_n + S\,y_n]\,, \text{ where} \tag{4.23}$$

$$A_{sq1} = \frac{\omega_{c1}\, M_{s0}\, dx^2}{\omega_{c0}\, M_{p0}}\,, \text{ and} \tag{4.24}$$

$$A_{sc1} = A_{sq1} + 3\,. \tag{4.25}$$

Equations 4.22 and 4.23 describe the first two rows of the transmission-line matrix that describes the cochlear model in Fig. 4.1. They account for a connection between the stimulus pressure and the pressure at the base of the cochlea.

4.2.3 The helicotrema boundary

The helicotrema boundary of the cochlea represents what happens with the remaining energy in the traveling wave when it reaches the apex. Ideally, all energy that was present in the stimulus should be absorbed when the apex of the cochlea is reached. To implement this, the last section of the transmission-line model could be represented by a matching impedance that has the same value as the impedance of the last section of the BM. A study by Puria and Allen (1991) has shown that there is no significant difference in the cochlear input impedance when representing the helicotrema with a short-cut or with a matching impedance. It was therefore chosen to represent the helicotrema boundary by a short-cut circuit, whereby little reflectivity from the helicotrema boundary was obtained. An expression for the pressure balance at the helicotrema was derived by applying Newton's second law to section n in the schematic of Fig. 4.1. Equation n in the transmission-line matrix is given by the following expression:

$$A_{hc}\, q_n - q_{n-1} = A_{sq}\, [D\, y_n + S\, y_n]\,, \text{ where} \tag{4.26}$$

$$A_{hc} = 1 + A_{sq} \tag{4.27}$$

and where A_{sq}, D and S were defined in Eq 4.14, 4.12 and 4.13 respectively.

4.2.4 Implementing tapering of the cochlea

To avoid low-frequency standing waves in the cochlear input impedance (Puria and Allen, 1991; Shera and Zweig, 1991), the reduced dimensions of the cochlea from base to apex (i.e., tapering) were considered in the model. This was accounted for by including the cochlear symmetry principle that suggested a proportionality between Y_p and Z_s (Zweig, 1991; Shera and Zweig, 1991). This proportionality was translated into:

$$C_p \propto M_s \,, \tag{4.28}$$

such that the BM compliance was made proportional to the acoustical mass of the cochlear fluids. Box-models of the cochlea that do not include tapering, assume that the acoustical mass of the cochlear fluids remains constant across the BM, whereas the compliance varies (de Boer, 1980; Viergever, 1980). From Zweig (1991), the following expressions for M_{sn} and C_{pn} were found:

$$i\omega\, M_{sn} = \frac{i\omega\, \omega_{c0}\, M_{s0}}{\omega_{cn}} \,, \text{and} \tag{4.29}$$

$$C_{pn} = C_{p0} \frac{\omega_{c0}}{\omega_{cn}} \,. \tag{4.30}$$

After applying (Zweig, 1991):

$$\omega_{cn} = \sqrt{\frac{K_{pn}}{M_{pn}}} \,, \text{and } C_{pn} = \frac{1}{K_{pn}} \,, \tag{4.31}$$

the following expressions were found for M_{sn} and M_{pn}:

$$M_{sn} = \frac{M_{s0}\, \omega_{c0}}{\omega_{cn}} \,, \text{and} \tag{4.32}$$

$$M_{pn} = \frac{K_0}{\omega_{c0}\, \omega_{cn}} \,. \tag{4.33}$$

Equations 4.32 and 4.33 were inserted into transmission-line equations 4.10, 4.23 and 4.26, which led to a transmission-line model of the cochlea that accounted for the reduced dimensions of the cochlea from base to apex.

4.2.5 The model parameters

The place-frequency distribution of the cochlea was chosen to agree with the experimental data of Greenwood (1961). The angular characteristic frequency was described by:

$$\omega_{cn} = 2\,\pi(f_{ch}\,10^{-\alpha\,n} - f_{cl})\,,\ \text{and} \tag{4.34}$$

$$\omega_{c0} = 2\,\pi\,(f_{ch})\,, \tag{4.35}$$

and f_{cl}, f_{ch} and α were defined in Table 4.1. From Zweig (1991) and Shera and Zweig

Table 4.1: Parameters used in the developed transmission-line model of the cochlea. The parameters (except for N) were based on the values chosen by van Netten and Duifhuis (1983), Duifhuis et al. (1986) and van Hengel (1996), and were chosen to represent the human cochlea.

Parameter	Value	Units	Physical Meaning
x	35e-3	m	BM length
b	1e-3	m	BM width
h	1e-3	m	Scala height
ρ	1e3	kg/m^3	Density cochlear fluid
f_{cl}	20682	Hz	High frequency limit
f_{ch}	140.6	Hz	Low frequency limit
α	61.765	-	
N	1.5	-	No. of wavelengths

(1991), the following expressions were found for the acoustical masses of the cochlear fluids and the BM at the base of the cochlea:

$$M_{s0} = \frac{2\,\rho}{b\,h}\ \text{with}\ [M_{s0}] = \left[\frac{M}{L^5}\right]\ ,\ \text{and} \tag{4.36}$$

$$M_{p0} = \frac{M_{s0}\,l^2}{(4\,N)^2}\ \text{with}\ [M_{p0}] = \left[\frac{M}{L^3}\right]\ ,\ \text{where} \tag{4.37}$$

$$l = \frac{1}{2.303\,\alpha}\,. \tag{4.38}$$

The stiffness constant K_0 was furthermore given by:

$$K_0 = M_{p0}\,\omega_{c0}^2\ \text{with}\ [K_0] = \left[\frac{M}{L^3T^2}\right]\,. \tag{4.39}$$

4.2.6 Methods for solving the transmission-line matrix

For every sample of stimulus p_{stim}, the transmission-line matrix consisting of equations 4.10, 4.23, 4.22 and 4.26, that were solved by using Gaussian elimination (Gentle, 1998). This allowed for an expression of $\dot{y}_n$ and y_n for every section of the BM, and thus a traveling wave. The model was solved over time by using the fourth order Runge-Kutta method. This method was adopted earlier by Diependaal et al. (1987) to solve transmission-line models over time. In this method, y and $\dot{y}$ of the previous time step were used to calculate the displacement at the next time-step Δt via $y_{t+\Delta t} = y_t + \dot{y}\Delta t$. At time 0, y and $\dot{y}$ were chosen to be zero.

4.2.7 Simulating reflection source otoacoustic emissions

CEOAEs and SFOAEs were modeled by including impedance irregularities along the length of the cochlea, such that coherent reflection can take place at the sites of these irregularities (Zweig and Shera, 1995). Random fluctuations were implemented on top of the damping constant δ such that D_n in the general transmission line equation 4.10 became:

$$D_n = \omega_{cn}\,\delta\,(1 + (0.02 * (rand_n - 0.5)))\,, \tag{4.40}$$

where $rand_n$ was a uniformly distributed number between 0 and 1. The random number generator was implemented in the model such that it depended on a starting "seed". Starting the generator with a specific seed always led to the same random number sequence across the BM. The choice of "test subject" in the model was linked to a specific seed such that different subjects, each with their own CEOAEs frequency pattern, could be simulated. Subject-dependent cochlear irregularities have been postulated to account for the frequency content of the CEOAE (Shera and Guinan, 1999).

4.2.8 Deriving the OAE

The pressure at the output of the model contained the middle-ear filtered stimulus and the pressure at the stapes. From Fig. 4.1, it follows that the output pressure was found

as:

$$p_{OUT} = p_{stim} + R_{ME}\, dx\, b\, \dot{y}_0 \tag{4.41}$$

When irregularities were added to the BM impedance, p_{OUT} contained a reverse traveling wave (i.e., the OAE) created by internal reflections of the forward traveling wave to these irregularities. To find the OAE in the model, two simulations were run: (i) a simulation where no BM irregularities were added and (ii) a simulation where cochlear irregularities were added such that an OAE was created. p_{OAE} was then found as the difference in output pressure obtained from these two simulations:

$$p_{OAE} = p_{OUTirr} - p_{OUTnoirr}\,. \tag{4.42}$$

4.3 Toward a nonlinear transmission-line model

The main purpose of the model was to simulate click-evoked otoacoustic emissions, thus the implemented nonlinearity was based on a study on animal BM impulse responses (Recio and Rhode, 2000). The data measured by Recio and Rhode (2000) in the left panel of Fig. 4.2 shows the near-invariance of the fine-structure in measured BM click responses recorded to increasing stimulus levels. Shera (2001b) developed a model in the frequency domain that could simulate this behavior through a careful placement of the poles of the describing BM admittance function $Y_p = 1/Z_p$:

$$Y_p(s) = s\, \left[\omega_{c0}\, M_{p0}\, \left(s^2 + \delta s + 1 + \rho \mathrm{e}^{-\psi s}\right)\right]^{-1}\,. \tag{4.43}$$

Zweig (1991) calculated the poles of Y_p and found two closely spaced poles near the real frequency axis of the complex plane, and an infinite series of poles further away from the real axis. The stability of the model depended on the two closely spaced poles that appeared near $f/f_{cf} = 1$, and their exact location was determined by the choice of δ, ρ and ψ. The feedback term in Eq. 4.43 was responsible for one of the two poles, and allowed for a model that was stable even when δ was chosen to be negative (i.e., active model). In another study, Shera (1992) and Zweig and Shera (1995) showed that the BM transfer function did not vary considerably when implementing a single coinciding pole rather than two closely spaced poles. Shera (2001b) used Eq. 4.43, and

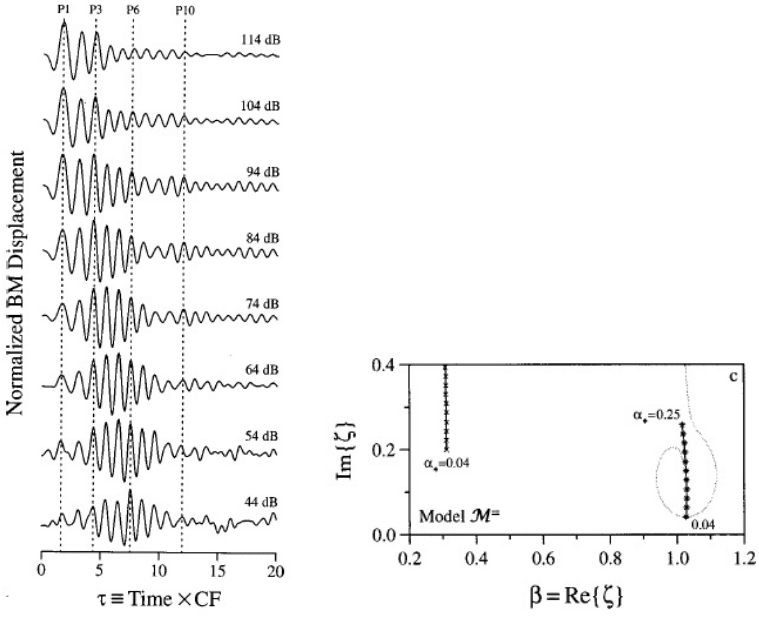


Figure 4.2: left: Near-invariance of fine time-structure in basilar membrane click responses. The figure shows normalized BM responses to clicks from Recio and Rhode (2000, Fig.2, chinchilla CB21). The horizontal axis measures time after the onset of umbo vibration in periods of the CF (14.5 kHz). Displacement waveforms are normalized to unit amplitude with peak click intensities indicated on the right in dB pSPL. Although the waveform envelopes vary systematically with click intensity, the underlying fine time-structure remains nearly invariant. right: BM admittance pole trajectories in the complex plane. ζ represents a scaling variable f/f_{cf}. Trajectories for increasing intensities were shown in the ζ plane of the principle frequency poles of the BM admittance. The pole trajectories were shown in function of α_*, the imaginary part of the double pole of the admittance. With increasing intensity, the double pole of the admittance $Y_p(\zeta;\alpha_*)$ moves along a curve nearly perpendicular to the real frequency axis from 0.04 to 0.25. Figures were reproduced from Shera(2001a) with permission from the author

chose the parameters δ, ρ and ψ such that a single coinciding pole at f/f_{cf} (or $\zeta = 1$) close to the real frequency axis was obtained. When the model was maximally active (i.e., δ < 0), the location of the imaginary part of the double pole of Y_p was located at $\alpha_* = 0.04$. The location of this double pole was indicated on the admittance pole trajectory in Fig. 4.2. Shera (2001b) designed the pole trajectory of Y_p to simulate the intensity-invariance of the temporal fine-structure of the impulse responses found by Recio and Rhode (2000). To account for the BM impulse responses in Fig. 4.2, the

pole trajectory moved from location $\alpha_* = 0.04$ at low stimulus intensities to location $\alpha_* = 0.25$ at high intensities.

The nonlinearity concept developed by Shera (2001b) that moves the location of the double pole of Y_p depending on local BM motion, and thus indirectly on the intensity of the input, was adapted and implemented in the present model. Firstly, the concept was extrapolated to all frequencies along the BM, and secondly, the concept was implemented such that the model determined the pole locations and thus Y_p adaptively. For a specific time step in the calculation, α_* was calculated depending on $\dot{y}_{BM}$ or y_{BM} of the previous time step. To implement an algorithm for this adaptive procedure, several steps were undertaken. Firstly, the influence of α_* on Y_p was investigated for different poles. Several linear versions of the model were tested for one fixed stimulus level, each with a different α_*. The simulated longitudinal $\dot{y}_{BM}$ and y_{BM} patterns were investigated to determine the range of α_* needed. Secondly, a link was made between the α_*, that gave rise to the desired longitudinal y_{BM} and $\dot{y}_{BM}$ patterns, and the intensity of the input. This was done by developing a function that described α_* in function of the $\dot{y}_{BM}$ and y_{BM} that corresponded to a specific stimulus level. The pole locations were chosen such that compressive behavior was obtained for stimulus levels between 30 and 80 dB SPL. Lastly, the algorithm was implemented such that for every time-step and for every location along the BM, α_* was calculated depending on the $\dot{y}_{BM}$ or y_{BM} of the previous time-step.

Several linear versions of the model were run for a 1-kHz sinusoidal stimulus with intensities between 0 and 100 dB SPL. These linear versions of the model only differed from each other in the location of the double pole α_*. The parameters δ, ρ and ψ were calculated by adopting the formulae derived in Shera (2001b) for different values of α_*:

$$a = \left(\alpha_* + \sqrt{\alpha_*^2 + c\,(1 - \alpha_*^2)}\right)/c\,,^{1} \tag{4.44}$$

$$\delta = 2\,(\alpha_* - a)\,, \tag{4.45}$$

[1] c = 120.899869

$$\psi = \frac{1}{2\,\pi\,a}\,,^{2} \tag{4.46}$$

$$\rho = \frac{\sqrt{1-\frac{\alpha_*^2}{4}}}{\frac{1}{2\,\pi\,a}\mathrm{e}^{1+\frac{\delta}{2\,\pi\,a}}}\,. \tag{4.47}$$

Fig. 4.3 shows displacement and velocity patterns across the length of the cochlea for stimulation with a fixed intensity 1-kHz sinusoid. The different curves represent simulations of different linear versions of the model where α_* was chosen to have a value between 0.04 and 0.4. δ, ρ, and ψ were calculated according to α_* via equations 4.45, 4.46, and 4.47, and were kept constant for all locations along the BM. The displacement and velocity patterns show the maximum values of y_{BM} and

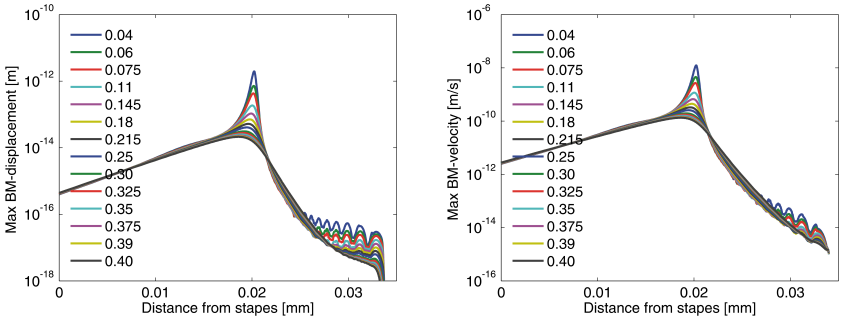


Figure 4.3: y_{BM} (left) and $\dot{y}_B M$ (right) patterns to a 1 kHz sinusoid with a fixed intensity. Every curve represents a simulation with a linear version of the model described by the pole location α_* that was indicated on the plot.

$\dot{y}_{BM}$ along the length of the BM. The patterns show tall and sharp peaks near CF when α_* was 0.04. For this pole location, the model was maximally active (i.e., $\delta <$ 0). When increasing the value of α_* to 0.4 (i.e., maximally passive model with $\delta > 0$), the patterns became broader, with a peak that reached its maximum more basally than for an α_* of 0.04.

To implement the nonlinearity, a link needed to be made between a specific α_* and

[2] Due to a memory constraint in the model implementation, ψ needed to be chosen constant for all α_*. A ψ value corresponding to α_*=0.06 was chosen, and preliminary simulations showed that the influence on Y_p of choosing a constant ψ was acceptable.

the intensity of the input. For low level inputs, it is desirable that the model behaves linearly and maximally active (i.e., small α_*), and for high input levels that it behaves passively and linearly (i.e., large α_*). For stimulus levels between 30 and 80 dB SPL, the model should behave compressively, and thus α_* should move from a more active value to a more passive value in this stimulus range. This idea was implemented by designing a pole-location function that either depended on y_{BM} or $\dot{y}_{BM}$, and indirectly on the intensity of the stimulus. Fig. 4.4 illustrates this concept by showing the levels of y_{BM} and $\dot{y}_{BM}$ at CF for a 1 kHz sinusoid with intensities between 0 and 100 dB SPL. The different curves show results for linear model implementations with

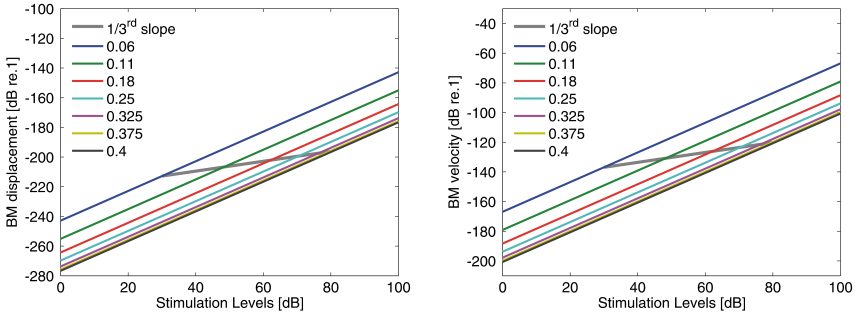


Figure 4.4: y_{BM} (left) and $\dot{y}_{BM}$ (right) for different stimulation levels. Each curve represents a simulation with a linear version of the model described by the pole location that was indicated on the plot. The desired nonlinearity was indicated by the gray line. For inputs below 30 dB active and linear behavior was desired; for levels between 30 and 80 dB, a compression factor of $1/3^{rd}$ was wanted. For higher stimulus levels, linear and passive model behavior was opted.

different α_* values. The gray line shows the desired compression factor of $1/3^{rd}$ for stimulation levels between 30 and 80 dB. At 30 dB, α_* was chosen to have a value of 0.06. The pole locations between 30 and 80 dB were chosen as the cross-section of the linear models with the $1/3^{rd}$ slope function. $\dot{y}_{BM}$ and y_{BM} were noted at the cross-sections for a specific stimulus level, together with α_*. Two pole-location functions were developed that related $\dot{y}_{BM}$ and y_{BM} to a specific pole, such that the desired nonlinearity function between 30 and 80 dB SPL was obtained, and this was illustrated in Fig. 4.5. The data points in Fig. 4.5 show the values of the pole locations in function of y_{BM} and $\dot{y}_{BM}$, taken from Fig. 4.4 such that a compression factor

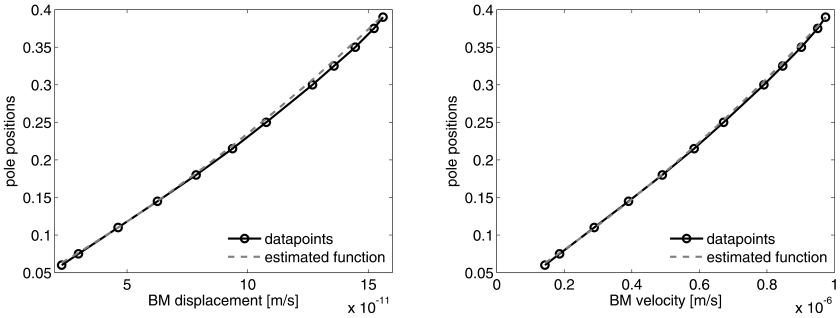


Figure 4.5: Location of the double pole α_* of Y_p in function of y_{BM} (left) or $\dot{y}_{BM}$ (right) to obtain a model that shows compressive behavior for stimulation levels between 30 and 80 dB SPL.

of $1/3^{rd}$ was obtained for stimulus levels between 30 and 80 dB SPL. A quadratic polynomial was fit to describe the pole locations PL as a function of y_{BM} and $\dot{y}_{BM}$:

$$PL_{\dot{y}} = 0.0023 \left(\frac{\dot{y}_{BM}}{\dot{y}_{TH}}\right)^2 + 0.0381 \left(\frac{\dot{y}_{BM}}{\dot{y}_{TH}}\right) + 0.0219\,, \tag{4.48}$$

and

$$PL_{y} = 0.0023 \left(\frac{y_{BM}}{y_{TH}}\right)^2 + 0.0381 \left(\frac{y_{BM}}{y_{TH}}\right) + 0.0219\,. \tag{4.49}$$

The pole-location functions in equations 4.48 and 4.49 were shown in Fig. 4.5, and determine the poles of Y_p in the model for a specific time step, depending on y_{BM} and $\dot{y}_{BM}$ of the previous time step. The implementation led to three regimes for the pole locations of Y_p, and these were summarized in Table 4.2. The y_{BM} and $\dot{y}_{BM}$ thresholds that were determined in Table 4.2, corresponded to sinusoidal stimulation at 30 and 80 dB SPL of a 1-kHz sinusoid at CF. The nonlinearity thresholds for the pole-location functions in equations 4.48 and 4.49 were also used to determine the poles of Y_p at other locations along the BM. This required for an extension of the nonlinearity thresholds for $\dot{y}_{BM}$ and y_{BM} in Table 4.2 to other locations along the BM. The transmission-line equations from Zweig (1991) imply a proportionality between $\dot{y}_{stapes}$ and $\dot{y}_n$. This means that if a flat spectrum $\dot{y}_{stapes}$ stimulus was applied at the base of the cochlea, the maximum velocity at each BM section would

Table 4.2: Pole locations PL that directly depend on BM vibration, and thus indirectly on stimulus intensity. Locations are given for a nonlinearity that tracks either $\dot{y}_{BM}$ (top) or y_{BM} (bottom)

Stimulus Level L	BM vibration	PL
L < 30 dB	$\dot{y}_{BM} < 1.429 \cdot 10^{-7}$	0.06
30 < L < 80 dB		Eq. 4.48 with $\dot{y}_{TH} = 1.429 \cdot 10^{-7}$
L > 80 dB	$\dot{y}_{BM} > 9.735 \cdot 10^{-7}$	0.39
L < 30 dB	$y_{BM} < 2.291 \cdot 10^{-11}$	0.06
30 < L < 80 dB		Eq. 4.49 with $y_{TH} = 2.291 \cdot 10^{-11}$
L > 80 dB	$y_{BM} > 1.561 \cdot 10^{-10}$	0.39

be identical. The nonlinearity thresholds that were determined for $\dot{y}_{BM}$ in Table 4.2 can thus be used for other locations along the BM. Due to the proportionality between $\dot{y}_{stapes}$ and $\dot{y}_{BM}$, the maximum y_{BM} at CF were proportional to $1/\omega_{cf}$. To account for this, the nonlinearity thresholds that were determined in Table 4.2 for y_{BM} were scaled by $1/\omega_{cn}$ in the model implementation.

Two types of nonlinearity strategies were included in the current model implementation: a nonlinearity that tracked $\dot{y}_{BM}$ and a nonlinearity that tracked y_{BM}. Both of these nonlinearities can be considered to act instantaneously on the local BM response. All poles along the BM, and thus the compression factors, at one time step were determined by the $\dot{y}_{BM}$ and y_{BM} pattern along the length of the BM of the previous time step. At time 0 in the model, $\dot{y}_{BM}$ and y_{BM} were set to 0, and all poles were chosen to be 0.06. For the next time step and onwards, the model determined the poles of Y_p adaptively depending on the $\dot{y}_{BM}$ and y_{BM} simulated to the next sample in the stimulus.

4.4 Evaluation of the Model

Rather than providing more details on the implementation of the model, this section shows key features of the model performance. The nonlinear adaptive transmission-line model was evaluated for several standard stimuli such as sinusoids and transients, and the results were compared to the relevant literature data. When transient stimulation was used, condensation pulses were filtered to account for the human middle-ear and ear-canal impedance as the current model implementation does not take these

impedances into account. Using a filtered transient as an input to the model would best represent the real situation where a condensation pulse was presented to the subject through an insert ear-phone. The filtered transient was a recording of an 80 μs click that was presented via the ER-2 loudspeaker to a BK-2012 ear-canal coupler, attached to the BK-4157 artificial ear. By presenting this filtered click to the model, a comparison between literature data and model simulations was made possible. In this section, longitudinal BM patterns were derived for sinusoidal and transient stimuli. Local BM transfer functions at the 1 kHz CF location were considered, and simulated frequency glides and auditory filters were evaluated.

4.4.1 Intensity dependence of longitudinal BM patterns and cochlear dispersion

Fig. 4.6 shows longitudinal y_{BM} (left panel) and $\dot{y}_{BM}$ (right panel) patterns in response to stimulation with a 1-kHz pure tone with intensities between 10 and 100 dB SPL. The ripples that can be observed in the y_{BM} and $\dot{y}_{BM}$ patterns near the

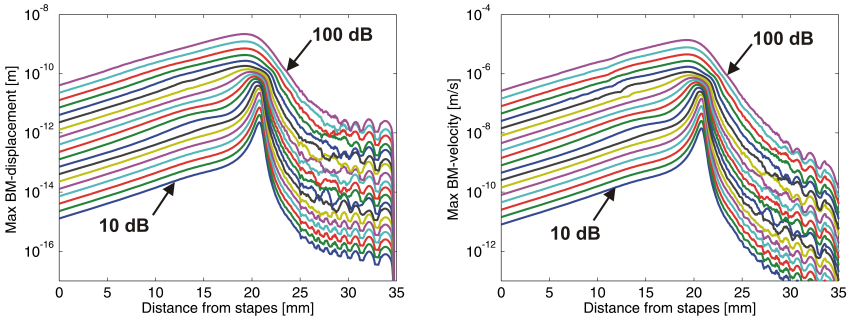


Figure 4.6: Longitudinal y_{BM} (left) and $\dot{y}_{BM}$ (right) functions for sinusoidal stimulation. The 1kHz stimuli ranged between 10 and 100 dB SPL in steps of 5 dB.

apex, were due to reflections of remaining energy in the model at the apical boundary of the model. The much slower and lower amplitude wiggles that can be observed on the basal site of the maxima of the patterns (e.g., in the mid intensity range of y_{BM}) represent reverse traveling waves or OAEs. Simulated BM compression was

investigated by deriving input/output functions at CF, and at one octave above and below CF. This was done for y_{BM} in Fig. 4.6, and the results were shown in the left panel of Fig. 4.7. At CF, BM compression with a slope of $1/3^{rd}$ was obtained. At

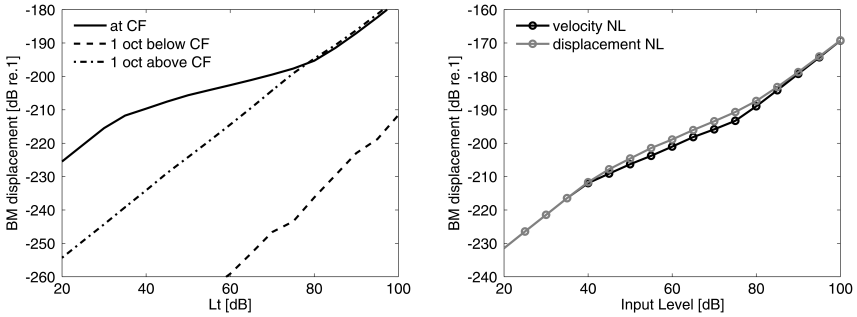


Figure 4.7: Left: input/output functions of y_{BM} simulated at 3 locations along the BM. The stimulus was a 1 kHz sinusoid that ranged between 10 and 100 dB SPL. The input/output function were measured at CF (1 kHz location) and an octave above and below CF. Right: input/output function for y_{BM} at the 1kHz location on the BM, measured to transient stimuli with levels between 20 and 100 dB peSPL. The nonlinearity was chosen to either track y_{BM} or $\dot{y}_{BM}$.

one octave above and below CF, linear behavior of the input/output functions was observed. The results in Fig. 4.7 agree qualitatively well with the BM input/output functions previously reported (e.g. Robles et al. (1986, Fig.2), Ruggero et al. (1997, Fig.6 and 7), Ren and Nuttall (2001, Fig.5) for sinusoidal stimulation). The right panel in Fig. 4.7 furthermore shows simulated input/output functions for transient stimulation, recorded at the 1-kHz CF location on the BM. Two different nonlinearity implementations were simulated. In the first, the nonlinearity tracked $\dot{y}_{BM}$ such that a compression factor of 0.4 was obtained for stimulation levels between 40 and 75 dB. In the second implementation, the nonlinearity tracked y_{BM}, which led to a compression factor of 0.5 for stimulation levels between 45 and 85 dB. Due to scaling symmetry in the model, the input/output functions for sinusoidal and transient stimulation would only look identical if a flat spectrum $\dot{y}_{stapes}$ was applied to the model. This was not the case for the transient stimulus used in Fig. 4.7, which explained the differences in compression observed in the input/output functions for sinusoidal and transient stimuli.

Other interesting level-dependent features of longitudinal BM patterns were observed in Fig. 4.6. The simulated y_{BM} and $\dot{y}_{BM}$ patterns became broader with increasing stimulus intensity and showed maxima that moved basally for higher intensities. As the α_* values in the model increased (i.e., more compression), the maxima of the longitudinal BM patterns were shown in Fig. 4.3 to move basally. This nonlinearity strategy resulted in y_{BM} and $\dot{y}_{BM}$ pattern maxima that moved basally for higher-intensity stimuli. The longitudinal BM patterns in Fig. 4.6 were in good agreement with longitudinal BM velocity patterns recorded by Rhode and Recio (2000, Fig.5) and Ren (2002, Fig.1) in chinchilla and gerbils, respectively. Both authors reported a basal shift of the pattern maxima, together with a broadening of the patterns for increasing intensities. Rhode and Recio (2000) linked the basal shift of the longitudinal patterns to the shift to lower frequencies of the maxima of local transfer functions. This can be verified in the simulations in Fig. 4.8 that show magnitude and phase transfer functions at the 1-kHz CF location for increasing intensities, this time for transient stimulation. For higher intensities, the magnitude-transfer functions

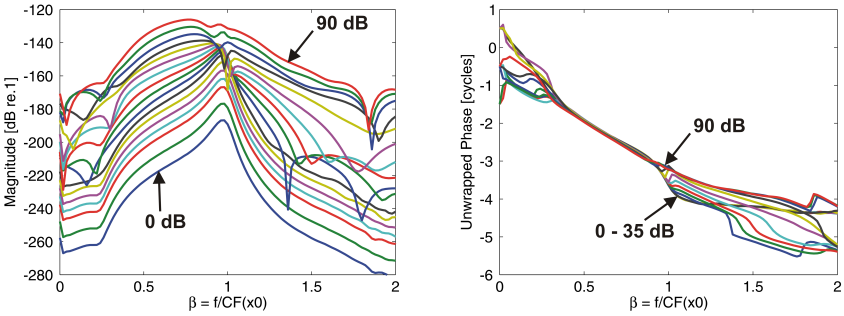


Figure 4.8: Left: Magnitude (left) and phase (right) transfer functions of y_{BM} at the 1-kHz location ($CF(x_0)$) on the BM. Transient stimuli with intensities between 0 and 90 dB peSPL in steps of 5 dB were used for the simulations. The transfer functions were plotted in function of the scaling variable β. Due to the scaling symmetry present in the model implementation, transfer functions with similar shapes (though different amplitudes) can be obtained for different CF.

moved their maxima (i.e., best frequency (BF)) to frequencies lower than CF. This shift to lower frequencies was also observed in Ren and Nuttall (2001, Fig.2) who reported a bit more than a half-octave shift in BF for the intensities of 10 and 100

dB SPL. They furthermore observed that the decrease in magnitude had a steeper slope for frequencies above BF than for frequencies below BF. The simulations in Fig. 4.8 agree with this finding for intensities below 75 dB SPL. The phase-transfer functions in the right panel of Fig. 4.8 show an increasing phase lead for higher stimulus intensities at CF. The phase-transfer functions measured by Ruggero et al. (1997, Fig.14) to pure tones showed similar phase leads at CF for higher-intensity stimuli. Thus, even though a half-octave shift in BF was observed in the local transfer functions in Fig. 4.8 for increasing intensities, the zero-crossings of the BM impulse responses remained intensity-invariant. The left panel of Fig. 4.9 demonstrates how a vertical shift of the dominant double pole α_* in Y_p for increasing intensities led to BM impulse responses at the 1-kHz CF location that showed increasing compression with increasing stimulus levels. At the same time, it can be observed from the left panel of

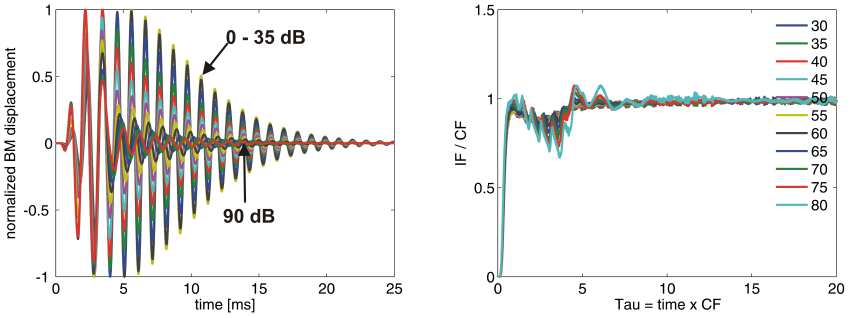


Figure 4.9: Left: y_{BM} at the 1-kHz CF location, for different intensities ranging between 20 and 90 dB in dB peSPL. The impulse responses were normalized by the pressure at the stapes of the cochlea. The nonlinearity strategy applied for these simulations tracked $\dot{y}_{BM}$. Right: Instantaneous frequency of y_{BM} at the 1-kHz CF location. The instantaneous frequency was calculated as the slope of the instantaneous phase pattern of the analytical signal of y_{BM}. The analytical signal was obtained by taking the Hilbert transform. The axes of the figure were made dimensionless, a technique that was earlier adopted by Shera (2001a) to compare frequency glides across frequency and species. Transient stimuli with intensities between 0 and 90 dB peSPL in steps of 5 dB were used for the simulations.

Fig. 4.9 that the temporal fine-structure of the BM impulse responses was the same for stimulus intensities below 80 dB. The simulated BM impulse response patterns were in good agreement with the BM impulse responses found in Recio and Rhode (2000)

(left panel of Fig. 4.2), except for levels above 80 dB[3]. The near-invariance of the instantaneous frequency at CF for increasing intensities is furthermore represented in the instantaneous frequency calculated at CF in the right panel of Fig. 4.9. The instantaneous frequency was calculated as the slope of the instantaneous phase pattern of the analytical signal of y_{BM}. It was observed from the right panel of Fig. 4.9 that at the 1-kHz location, the instantaneous frequency moved to CF after about 5 ms. This phenomenon has been named the "frequency glide", as it takes some time for a section of the BM to oscillate at its CF. The modeled results in Fig. 4.9 show good agreement with the experimental data of Shera (2001a, Fig.2 and 3). Frequency glides have been shown to be a consequence of traveling wave dispersion and reflect a time dependence of the effective driving force at the measurement location (Shera, 2001a). In the frequency domain, cochlear dispersion can be seen as the increase in group delay of the transfer function as the stimulus frequency approaches CF. Fig. 4.10 shows an increase in the group delay of the transfer function measured at the 1-kHz location. These results agree qualitatively well with the group delay decrease observed at CF in Ruggero et al. (1997) for increased stimulus intensities. It can also be observed from Fig. 4.10 that the group delay abruptly decreases to negative values at CF. This negative group delay may be the result of cochlear reflections that occur at CF.

As a consequence of local scaling symmetry, transfer functions measured as a function of frequency at a fixed BM location, also describe the traveling wave as a function of BM location at a fixed frequency (Zweig, 1976). Applying this principle to the transfer functions in Fig. 4.10 and 4.8, led to the 1-kHz traveling waves in Fig. 4.11. Due to scaling symmetry, there is also a correspondence between glides in the frequency domain, and in the wavelength of the traveling wave in the spatial domain. Glides can be seen as the decrease in wavelength of the traveling wave as CF is approached (Shera, 2001a). This decrease in wavelength near CF can be observed for the 1-kHz traveling waves in Fig. 4.11. From this figure, it can be observed that there were about 1.5 cycles in the traveling wave before CF was reached. Cochlear

[3] The lag in the BM impulse responses observed for intensities above 80 dB, was probably due to the model constraint of a constant ψ in Y_p for all α_*. The effect of keeping ψ constant would be most pronounced for the higher stimulus intensities.

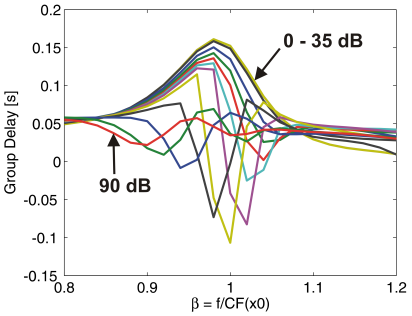


Figure 4.10: Group delay of y_{BM} at the 1-kHz location on the BM. Transient stimuli with intensities between 0 and 90 dB peSPL in steps of 5 dB were used for the simulations. The group delay was calculated as the negative slope of the phase transfer function at CF.

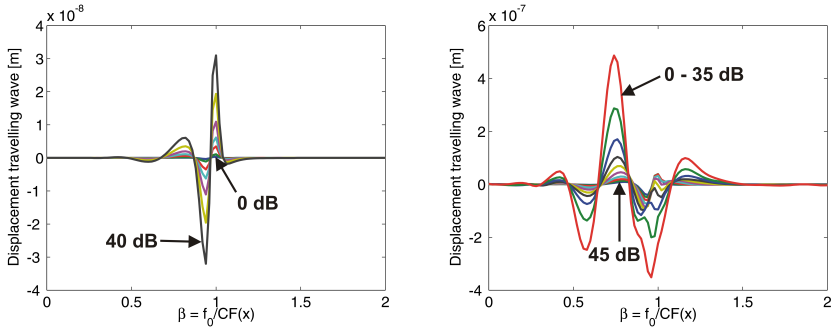


Figure 4.11: Snapshot of the 1-kHz traveling wave of y_{BM} for intensities between 0 and 40 dB peSPL (left) and 45 and 90 dB peSPL (right). Transient stimuli with intensities between in steps of 5 dB were used for the simulations. The scaling symmetry principle allows for a calculation of the traveling wave from the magnitude-transfer function via a transformation of the x-axis from $\beta = f/CF(x_0)$ to $\beta = f0/CF(x)$

dispersion in the model was clearly demonstrated by the simulated glides and transfer functions in Fig. 4.9, 4.10, 4.11.

Still, there was one missing link between the intensity-invariance of the glides in Fig. 4.9 (and the BM impulse response zero-crossings) and the half-octave shift in BF that occurred for the magnitude transfer functions in Fig. 4.8. Shera (2001a) and Shera (2001b) linked the two properties together by stating that the resonance frequencies

of Y_p (and thus cochlear dispersion) do not show a stimulus-intensity dependence. The half-octave shift, on the other hand, was due to global properties (i.e., mechanical properties of the BM) of the driving pressure and would therefore show an intensity dependence (Shera, 2001b). Carney et al. (1999) found a similar explanation for the half-octave shift, and both studies explained the effect as the "centre of energy" of the pressure wave at CF that occurs earlier for higher intensities. As a consequence, the pressure maxima would occur within the glide period of that section of the BM. As within glide period, the section of the BM usually responds to lower frequencies than CF, this could lead to a CF location that is dominated by driving frequencies lower than CF. As a consequence, a shift in BF to lower frequencies would be observed for these higher intensity stimuli (Shera, 2001b).

4.4.2 Scaling Symmetry

As a consequence of the scaling symmetry that was adopted in the describing transmission-line equations 4.5 and 4.6, $\dot{y}_n$ scales with $\dot{y}_{stapes}$. This scaling property was tested in Fig. 4.12 for transient stimulation. $\dot{y}_{stapes}$ was calculated in accordance to the schematic in Fig. 4.1 as:

$$\dot{y}_{stapes} = \frac{2 * (p_0 - p_1)}{b\,dx\,M_{s0}}\,, \tag{4.50}$$

and its magnitude spectrum was plotted in Fig. 4.12. The maxima of the simulated

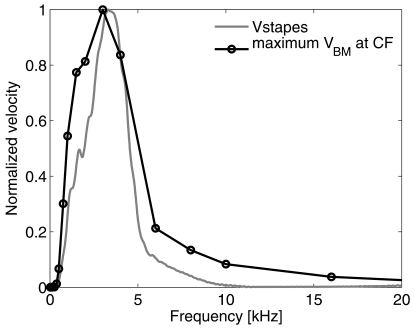


Figure 4.12: Normalized $\dot{y}_{stapes}$ and maximal $\dot{y}_{BM}$ simulated at different locations along the BM for transient stimulation.

$\dot{y}_n$ impulse responses measured at different CF locations in Fig. 4.9, were plotted in Fig. 4.12. A good agreement was found between the maxima of $\dot{y}_n$ measured at different CFs, and $\dot{y}_{stapes}$ and it can thus be concluded that scaling symmetry was represented well in the model implementation. The magnitude spectrum of $\dot{y}_{stapes}$ in response to the transient stimulus contained mainly energy in the 1–5 kHz region. This was a consequence of filtering the condensation pulse with a middle-ear and ear-canal impedance before it was applied to the model.

4.4.3 Auditory filters

The left panel in Fig. 4.13 shows simulated y_n impulse responses measured at different locations on the BM, for stimulation with a transient of 60 dB peSPL. Shorter impulse

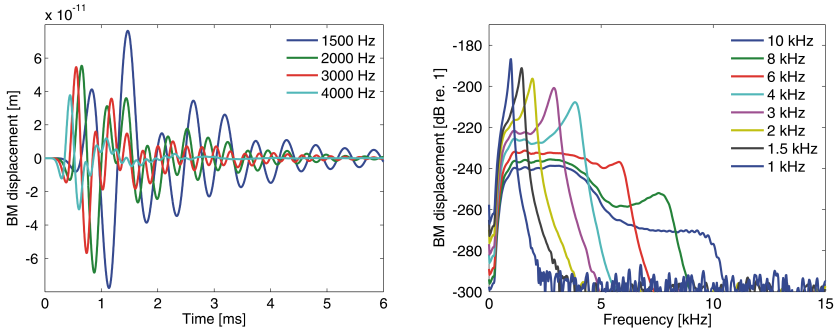


Figure 4.13: Left: y_{BM} at different locations along the cochlear partition. The stimulus was a transient with level of 60 dB peSPL. Right: Magnitude spectra of y_{BM} measured at different locations on the BM (i.e., auditory filters). The stimulus was a transient.

responses were obtained at higher CF locations than at lower CF locations. This was in good agreement with the impulse-response data presented in Recio and Rhode (2000, Fig.7). The different resonance frequencies observed in the left panel in Fig. 4.13 for different sections of the BM, can be interpreted as different auditory filters. The right panel in Fig. 4.13 shows magnitude-transfer functions of y_{BM} measured at different locations along the BM. Only eight auditory filters were plotted in the right panel of Fig. 4.13; in reality the cochlear model consisted of as many auditory filters as there were sections in the BM (i.e., 2000 in this implementation). The magnitude-transfer

functions showed maxima at their CF, except for the filters measured at the 6, 8 and 10 kHz CF. These three filters show peaks around CF, but also contain energy in the mid-frequency range. This behavior can be addressed to the stimulus that corresponded to a $\dot{y}_{stapes}$ magnitude spectra that had its energy centered in the 1 to 5 kHz range (see Fig. 4.12). The auditory filters were seen to become broader with increasing center frequency, a feature that has been described before both in BM tuning curves (e.g. Robles et al. (1986)), and in psychophysical experiments (e.g. in Glasberg and Moore (1990)). The damping constant δ was kept constant in the present model, leading to auditory filters that had a constant quality factor $Q = 1/\delta$. As a consequence, the bandwidth of the filter doubled for every doubling of CF.

Keeping a constant Q across the BM allows for applying local scaling symmetry in the modeled cochlea. Several other time-domain models of the cochlea have for a variable Q along the length of the BM (e.g. Talmadge et al., 1998; van Hengel, 1996), and thereby broke the local scaling symmetry principle. Since there are currently conflicting theories about the different OAE-measures that claim to reflect the width of the auditory filters across frequency (Shera et al., 2002; Keefe et. al, 2007; Lineton and Wildgoose, 2009), the benefit of modeling human auditory filters with a varying Q across frequency is uncertain. Local scaling symmetry has been observed in both BM transfer functions (e.g. Rhode, 1971; Rhode and Cooper, 1996), and neural tuning curves (e.g. Liberman , 1978), thus this principle was preferred over a variable Q in this model implementation. Even though scaling symmetry breaks down near the apex due to the uncertainty about the cochlear place-frequency map for low frequencies, the concept has been shown powerful in describing concepts such as traveling wave propagation and gain (Shera, 2007), and cochlear dispersion (Shera, 2001a).

4.5 Summary

A transmission-line model of the cochlea was developed in this chapter based on the transmission-line equation by Zweig (1991) that described the mechanics of a section of the cochlea as a function of the scaling variable $f/f_{cf}(x)$. The transmission-line equation by Zweig (1991) was applied to a model that divided the cochlea into 2000 sections. Two boundary conditions were furthermore developed that described the

connection of the cochlea with the middle ear and the helicotrema. The developed transmission-line matrix was solved such that an expression for the traveling wave along the length of the cochlea was obtained. The model was furthermore solved over time by using the fourth order Runge-Kutta method, such that a traveling wave was obtained for every sample in the stimulus. The model was adjusted to simulate reflection-source emissions by including impedance irregularities along the length of the cochlea. Distortion-source emissions and BM compression were simulated by including a nonlinearity in the model. The model employed a nonlinearity strategy that was based on a description of the nonlinear behavior of recorded BM impulse responses (Shera, 2001a).

The developed model accounted for an intensity invariance of the zero-crossings of the BM impulse responses, in agreement with the data of Recio and Rhode (2000). At the same time, the model simulations were able to demonstrate key features of cochlear dispersion such as frequency glides. The intensity-invariance of the frequency glides was furthermore demonstrated to be in good agreement with the data in Shera (2001a). Lastly, a half-octave shift in the maxima of the longitudinal BM patterns was observed to correspond well to experimental data by Rhode and Recio (2000) and Ren (2002).

Chapter 5 employs the transmission-line model developed in this chapter to investigate level-dependent features of CEOAEs. Firstly, the model performance in simulating CEOAEs is investigated, by evaluating the simulated CEOAE level-curve and group delay. Secondly, the cochlear origin of temporal adaptation in the CEOAE level-curve is investigated. More specifically, the question whether dynamic changes in the BM nonlinearities are necessary to account for dynamic changes in the CEOAE level-curve is addressed.

5

On the Origin of Temporal Adaptation of the CEOAE Level-Curve [1]

This study investigates the cochlear origin of temporal adaptation in the click-evoked otoacoustic emission (CEOAE) level-curve. The dynamics observed in the CEOAE level-curve when a suppressor click is presented 0–10 ms before a test click, was linked to time-dependent changes in local BM input/output functions. This was established by using a time-domain model of the cochlea that simulates level-dependent features of the CEOAE. The model was implemented with a time-invariant nonlinearity to test the hypothesis that temporal overlap of local BM impulse responses would account for temporal adaptation of the CEOAE level-curve. Experimental and simulated results were compared, and revealed that the model was able to simulate temporal adaptation well for the linear and knee-point region of the CEOAE level-curve. However, the model simulations did not account for the sustained suppression levels observed in the compressive region of the experimental CEOAE level-curves. The presented model, implemented with a time-invariant nonlinearity, was thus unable to account for important features of temporal adaptation in the CEOAE level-curve. These results suggest that a mechanism other than a time-invariant nonlinearity that tracks the BM impulse response fine-structure, underlies temporal adaptation.

5.1 Introduction

Temporal adaptation of the click-evoked otoacoustic emission level-curve is the change in the level-curve that can be observed when a suppressor is presented 0–10 ms before

[1] This chapter was presented on the 2010 Midwinter Research Meeting of the Association for Research in Otolaryngology in Anaheim, CA

a test click. The experimental data in Chapter 3 suggested that the origin of temporal adaptation was found in local place-fixed BM mechanisms. The exact nature of the dynamics involved in the BM mechanisms was not found, but two hypotheses were put forward. In the first, the nonlinearity was assumed to be static or time invariant (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Hine and Thornton, 2002; Harte et al., 2005). Temporal adaptation would in this framework be caused by a constructive summing of local BM impulse response patterns. The duration of these impulse response patterns (Recio and Rhode, 2000) causes them to overlap in time when two clicks are presented closely in time to each other. If constructive summing takes place within this overlap period, then the input to the nonlinearity would be larger than for a single click. As a result, more compression would be applied by the static nonlinearity, which would lead to increased suppression levels. In the second hypothesis, a time dependence that operates on a time scale of 0–10 ms would exist in the local BM nonlinearities (Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Verhulst et al., 2008). In this framework, the BM input/output function would dynamically shift horizontally for a duration of 0–10 ms after a click is applied. When a second click is presented within this time period, it would operate from a different BM input/output function than the first click, whereby different compression characteristics would be obtained. This study aims at investigating the origin of temporal adaptation by using a time-domain model of the cochlea that simulates otoacoustic emissions. The details of this transmission-line model of the cochlea were thoroughly discussed in chapter 4, where it was demonstrated that the model was able to simulate local BM impulse responses and I/O functions in agreement with recorded data of Recio and Rhode (2000), Rhode and Recio (2000) and Ren (2002).

Using a model approach to investigate the origin of temporal adaptation, allows for an investigation of the relation between simulated BM I/O functions and CEOAE level-curves. As the CEOAE level-curve contains information about several locations along the BM, it reflects a broadband level-function. This study investigates the relation between the CEOAE level-curve and local BM I/O functions, when two clicks were presented closely in time to each other. Firstly, it was verified how well the model performed in simulating key features of CEOAEs, such as the level-curve and group

delay. Secondly, the hypothesis stating that temporal overlap of local BM impulse responses underlies temporal adaptation was tested. For this purpose, the model was implemented with an instantaneous nonlinearity, meaning that the BM compression properties at one time instant changed depending on the BM velocity of the previous time step. The model did not include time-dependent processes in the BM nonlinearities, and did thus not model adaptation in OHC motility. If temporal adaptation in CEOAEs would be adequately modeled with an instantaneous nonlinearity, then temporal overlap of BM impulse responses could explain the effect. A rejection of the hypothesis would strongly suggest that an additional mechanism, such as outer hair cell adaptation, would underly temporal adaptation in CEOAEs.

5.2 Materials and Methods

To allow a comparison between measured data and model simulations, the model should describe the human system adequately. The aim was to use the original experimental setup that was laid out in paragraph 3.2 of Chapter 3 for the model simulations. Several adaptations were made to allow for a comparison between recorded and simulated CEOAEs. Firstly, the transmission-line model of the cochlea developed in Chapter 4 did not include a human ear-canal and middle-ear impedance. In the human CEOAE experiment, an 80 μs click was applied through an ER-2 probe that was placed in the ear canal of the subject. The pressure at the base of the human cochlea was thus influenced by the ear-canal and middle-ear impedance. To account for this in the model, the 80 μs click was filtered by using an ear-canal and artificial-ear impedance. The 80 μs click was presented via the ER-2 loudspeaker to a BK-2012 ear-canal coupler, attached to the BK-4157 artificial ear, and the recording was normalized before it was applied as an input to the model. The resulting stimulus is shown in Fig. 5.1. The normalized transient was amplified in the model such that its level (peSPL) corresponded to the desired stimulus level. It can be observed from Fig. 5.1 that the transient stimulus had a broadband spectrum with energy concentrated in the 0.5 to 5 kHz region. As demonstrated earlier in Fig. 4.12, the velocity at the base of the cochlear model showed a similar spectrum as in Fig. 5.1 to this transient stimulus.

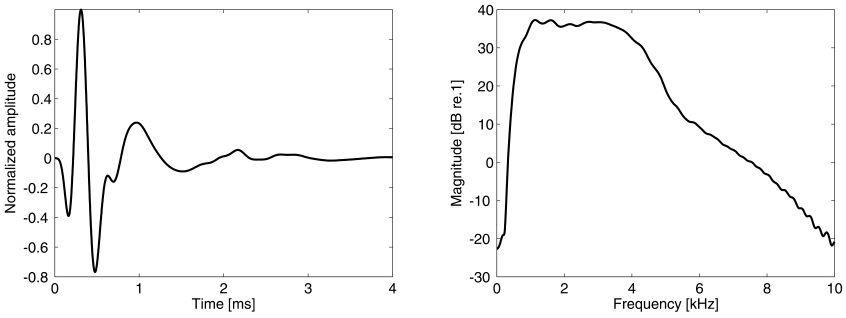


Figure 5.1: Left: Normalized recording of an 80 μs condensation click in a BK-2012 ear-canal coupler that was attached to the BK-4157 artificial ear. Right: Magnitude spectrum of the normalized transient stimulus.

The simulated CEOAEs arose through BM irregularities that were placed randomly in the 0.5 to 20 kHz region of the cochlear model. As the stimulus contained energy in the 0.5 to 5 kHz region, most reflections were expected to come from this frequency region. In the CEOAE experiment, the recorded OAE was band-pass filtered between 0.5 and 5 kHz for noise reduction purposes. The frequency content of the recorded and simulated OAE were thus expected to show a similar broadband spectrum. The results in Chapter 3 showed that all four subjects showed temporal adaptation, regardless of the bandwidth of their CEOAE spectrum. Thus, a model that simulates a broadband CEOAE spectrum is well suited to investigate temporal adaptation. The simulated results were compared to the data obtained for subject #1 in chapter 3.

5.2.1 Experimental paradigm

The experimental paradigm presented in Chapter 3 for the derivation of the unsuppressed (US) and derived suppressed (DS) response, was also adopted for the model simulations. As indicated in Fig. 3.2, two transient stimuli were presented to the model. In the first (DC), two condensation transients (i.e., the suppressor and test click) were presented with a specific inter-click interval (ICI) between them. In the second stimulus (DCI), a condensation and rarefaction transient, separated by the

same ICI, were presented. The derived suppressed response (DS) was obtained by subtracting the DCI response from the DC response and by halving the result. The unsuppressed response (US) was obtained by presenting a condensation transient with the level of the test click to the model. The two experiments conducted in Chapter 3 were simulated with the model. In the first experiment, the time scale of temporal suppression was investigated by varying the ICI between 0 and 8 ms for test click levels (L_t) of 40, 50, 60 and 65 dB peSPL. The suppressor level L_s was first chosen to equal L_t, and later to be 3 and 6 dB lower. In the second experiment, unsuppressed and suppressed level-curves were derived for the ICIs of 1, 1.5 and 2 ms. L_s and L_t had equal levels and ranged between 20 and 90 dB peSPL in 5 dB steps.

5.2.2 Post-processing

For the human recordings, it was necessary to separate the CEOAE from the early-latency linear reflection that was present in the recordings. The OAE was set to start 6 ms after click onset because the linear reflection due to the middle ear and ear canal had not decayed down to the noise floor before that point in time. In the model simulations, another technique was adopted to separate the OAE from the linear impulse response. Two simulations were made for every ICI and level configuration. In the first simulation, the BM irregularities in Eq 4.40 were set to zero such that the resulting output pressure at the base of the cochlea (i.e., p_{OUT} in schematic 4.1) showed a response determined by the stimulus and the cochlear load. In the second simulation, BM irregularities were included in the cochlear impedance function, such that a reverse traveling wave was present in the output pressure at the base of the cochlea. As explained in section 4.2.8, a subtraction of both responses led to the CEOAE. In contrast to the experimental separation technique, the model simulations allowed for an analysis of the early-latency region (i.e., 0–6 ms) of the OAE.

The output pressures of the simulated and recorded CEOAEs were not comparable in their amplitude. This was in part due to the choice of the parameters that describe the cochlear model. The output pressure of the model was furthermore measured at the stapes (see Fig. 4.1) and did not travel through the middle ear and ear canal as for human subjects. Hence, the pressure at the stapes was different than what it would be at the ear-drum. Conversely, the waveform shape and intensity-

dependence of the simulated waveforms were comparable to the recorded CEOAEs. Suppression was calculated as the level difference between the unsuppressed and suppressed CEOAE via:

$$\text{Suppression}(t_p) = L_{US}(t_p) - L_{DS}(t_p)\,, \text{ where} \tag{5.1}$$

$$L(t_p) = 20\log_{10}\left[\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2\right]^{1/2}, \tag{5.2}$$

where N depicts the number of samples in the analysis window, x_i the i^{th} sample of the CEOAE, and μ_x the mean of the CEOAE within the considered window t_p. Even though the exact amplitudes of the simulated and recorded CEOAE were not comparable, suppression levels were. Suppression levels were calculated between $\tau_s = 6$ and 18 ms for the short-latency CEOAE, and between $\tau_s = 24$ and 36 ms for the long-latency CEOAE. For the model simulations, a third early-latency window was also considered between $\tau_s = 0$ and 6 ms.

5.3 Results

Fig. 5.2 shows the waveforms of the recorded (top) and simulated (bottom) CEOAE. The most pronounced difference between the two CEOAEs was found in the amplitude reduction with latency. The recorded CEOAE showed long-latency components (i.e., SSOAEs) that were ringing with a steady amplitude until the end of the recording window. The simulated CEOAE, on the other hand, showed a pronounced amplitude reduction with latency. Still, some low level long-latency energy was also present in the simulated CEOAE. The differences observed for the long-latency region of the CEOAEs were also reflected in the frequency spectra as shown in Fig. 5.3. The long-latency (LL) spectra of subject #1 showed energy in the 0.5–2 kHz region and in the 4–5 kHz region, whereas the model only contained frequency components in the 1–1.5 kHz region. The simulated and recorded short-latency (SL) spectra were in better agreement. The modeled CEOAE showed a broadband spectrum with energy between

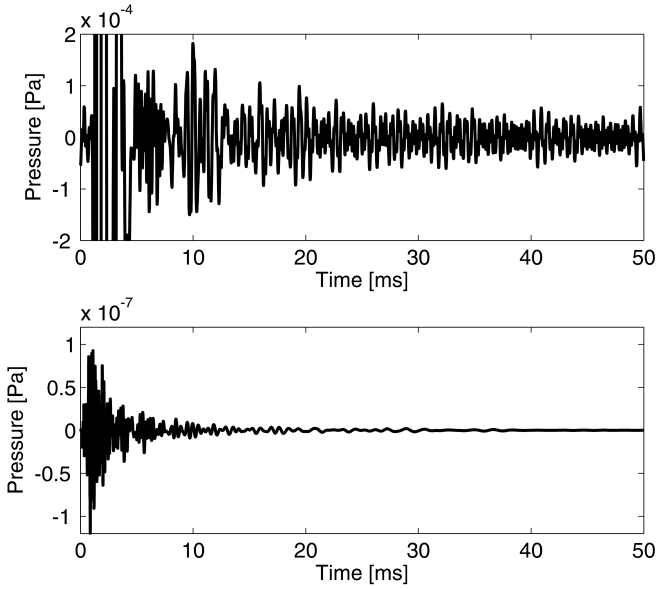


Figure 5.2: Top: Click response recorded in the ear-canal of subject #1 to an 80 μs condensation click of 65 dB peSPL. The response contains the early-latency linear response (i.e., 0–6 ms) and the CEOAE (> 6 ms). Bottom: Simulated CEOAE for the transient stimulus in Fig. 5.1 at 65 dB peSPL. Note that whereas the reflected click was still present for the subject (i.e. 0–4ms), the linear click response was removed from the simulated CEOAE.

0.5 and 5 kHz, and the recorded CEOAE contained energy in the same region, except for the 2–3.5 kHz region.

The short- and long-latency spectra for the recorded CEOAE showed a very similar frequency content due to the presence of SOAEs. SOAEs, that are thought to arise as standing waves in the cochlea due to multiple internal reflections of energy (Shera, 2003), were not modeled adequately. A model that accounts for SOAEs, should have a good reflectivity at the base, such that standing waves can occur. As observed from the long-latency frequency spectrum in Fig. 5.3, the model only showed good reflectivity (and thus energy in the long-latency region of the CEOAE) in the 1–1.5 kHz frequency range. The reason for this low reflectivity was due to the matching impedance implemented at the base of the cochlea that absorbed most of the energy in the reverse traveling wave. The results in chapter 3 showed that both purely

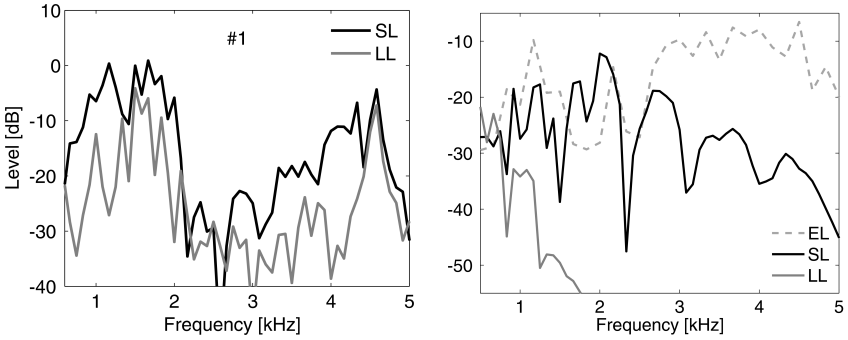


Figure 5.3: Frequency spectra of the recorded (left) and simulated (right) CEOAE for different time windows in the response: early-latency (EL; 0–6 ms), short-latency (SL; 6–18 ms) and long-latency (LL; 24–36 ms).

evoked and spontaneous emissions were similarly affected by temporal adaptation. The lack of well pronounced SOAEs in the presented model would thus not make it less suitable for investigating temporal adaptation. The spectral differences in the short- and long-latency regions of the simulated CEOAE can furthermore be used to investigate whether temporal adaptation was different for different frequency regions. Even though there were clearly pronounced differences in the waveform and spectra of the modeled and simulated CEOAE, it is assumed that the model is well suited to investigate the main hypothesis proposed in this study. If temporal adaptation can be shown to work in a model implemented with an instantaneous nonlinearity and a broadband frequency spectrum, then subject- or frequency-dependent features of the effect can be investigated more thoroughly with a series of tailored models. Tailored models can be implemented through a careful placement of the BM irregularities, such that the frequency content of the simulated and recorded emission would correspond closely.

5.3.1 Level-dependent features of the CEOAE

Fig. 5.4 shows normalized waveforms of the recorded (left panel) and simulated (right panel) CEOAE as a function of stimulation level. Both waveforms showed increased

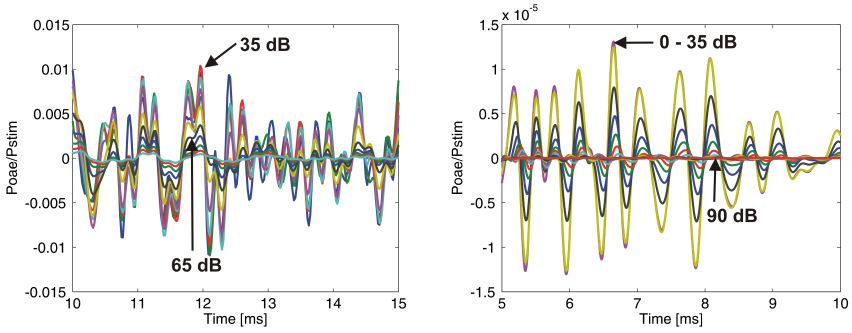


Figure 5.4: Left: Waveforms of the recorded CEOAE to a condensation click with stimulus intensity between 30 and 85 dB peSPL in 5 dB steps. Right: Waveforms of the simulated CEOAE to a transient with stimulus intensity between 0 and 90 dB peSPL in 5 dB steps. The waveforms were normalized to the pressure of the stimulus.

compression, observed as a reduction in the normalized CEOAE waveform, for increased stimulus intensities. The level-dependent compression observed in the CEOAE waveforms was furthermore illustrated in Fig. 5.5, that shows level-curves for the recorded (left panel) and simulated (right panel) CEOAE. Whereas the recorded CEOAE

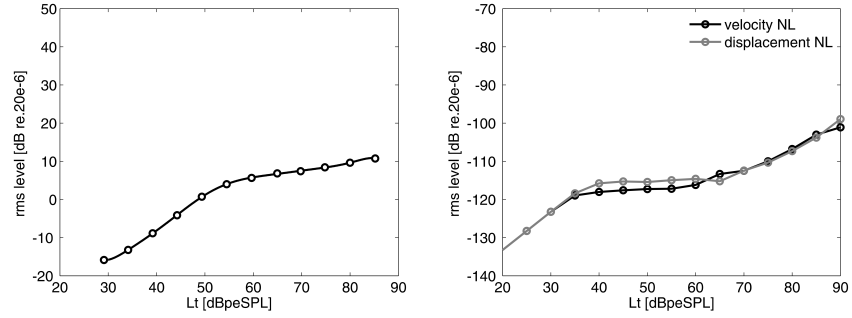


Figure 5.5: CEOAE level-curves of the recorded (left) and simulated (right) CEOAE. Left: rms levels were calculated for the 6–36 ms time window of the CEOAE. Right: rms levels were calculated for the whole CEOAE waveform (0–36 ms). Two nonlinearity strategies were tested: in the first the instantaneous nonlinearity tracked the BM velocity, in the second the BM displacement was tracked.

showed compressive behavior for stimulus intensities between 50 and 85 dB peSPL,

the simulated CEOAE showed compression for intensities between 35 and 65 dB peSPL. The simulated CEOAE level-curve furthermore grew more linearly for stimulus levels above 70 dB peSPL. Within the compressive region of the level-curves, compression factors of 0.27 dB/dB and 0.03 dB/dB (displacement NL) and 0.04 dB/dB (velocity NL) were obtained for the recorded and simulated CEOAE respectively. The lower compression threshold for the model simulations was partly due to the nonlinearity threshold that was set at 30 dB in Eq. 4.48 and 4.49. It is furthermore possible that the stimulus presented to the subject was attenuated due to the travel through the middle ear (Puria, 2003), such that the stimulus pressure at the stapes was lower than that assumed in the model. To account for the difference in knee-point location of the recorded and simulated CEOAE level-curve, lower stimulation levels were used for the model simulations in experiment 1. From the simulated CEOAE in Fig. 5.5 it was furthermore observed that both nonlinearity implementations (i.e., nonlinearity tracking y_{BM} or $\dot{y}_{BM}$), showed qualitatively similar level-curves. The following simulations were therefore only carried out for a model that was implemented with an instantaneous nonlinearity that tracked $\dot{y}_{BM}$.

The zero-crossings of the recorded CEOAE in Fig. 5.4 were found to remain nearly at the same location for different intensities. In agreement with the observations of Carvalho et al. (2003), small latency differences in the peaks of the recorded waveforms were observed for higher stimulus intensities. No systematic pattern was observed as to when a lead or lag occurred, similar to what was observed in the CEOAE waveforms of Carvalho et al. (2003). The modeled CEOAE showed fixed zero-crossings for stimulus intensities up to 55 dB peSPL, whereas latency differences in the waveform fine-structure were observed for higher intensities. These latency differences were quantified by investigating the group delay of the simulated CEOAEs in Fig. 5.6. Group delays were calculated as the negative slope of the unwrapped phase spectrum of the simulated CEOAEs of 40 different "subjects". Each subject represented a model implementation for a different random series of BM irregularities, which automatically led to broadband magnitude spectra that contained different frequency components. The mean group delay across subjects was shown in the left panel of Fig. 5.6 for a stimulus level of 50 dB. At 1 kHz, a mean group delay of 12.5 ms was found, and this value decreased with increasing frequency to 7 and 5 ms at 2

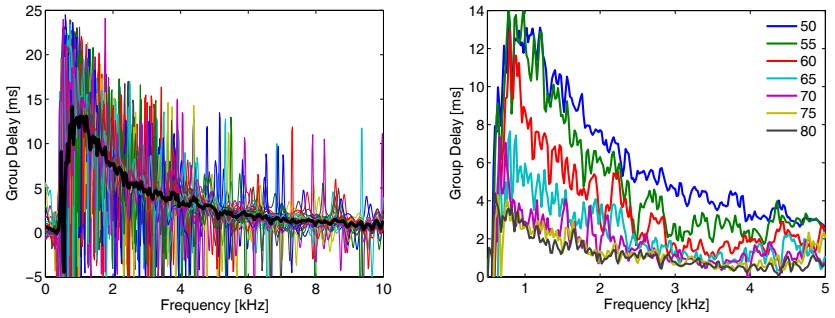


Figure 5.6: Left: Group delay of the simulated CEOAE for 40 different "subjects", each with a different random placement of the BM irregularities along the BM. A transient stimulus of 50 dB peSPL was used and the thick black curve shows the mean group delay. Right: Mean group delay of the simulated CEOAEs of 40 subjects for different click intensities.

and 4 kHz respectively. These values were within the range of the experimental group delays found by Pigasse (2008) for tone-burst OAEs at 66 dB peSPL (i.e., 9.5, 7 and 4 ms at 1, 2 and 4 kHz), and Shera et al. (2002) for low level SFOAEs (i.e., 11, 7 and 4.5 ms at 1, 2 and 4 kHz). The mean group delays for different stimulus intensities in the right panel of Fig. 5.6 show a decrease in overall group delay with increasing stimulus intensity. These results imply that, as the stimulus intensity increased, the frequency components in the CEOAE waveforms in Fig. 5.4 occurred at earlier times. For the levels of 50 and 55 dB, there was little change in the group delay in the 1–1.5 kHz region, thus the zero-crossings for these levels were similar in the time frame visualized in Fig. 5.4. Schairer et al. (2006) calculated the 1 kHz SFOAE group delay for different intensities and found values of 9.5, 8, 7 and 6 ms for stimulus levels of 50, 60, 65 and 70 dB. At 1 kHz, the model simulations in Fig. 5.6 also showed a decrease in group delay with values of 12.5, 9, 6 and 3 ms for the same stimulus intensities. The simulated group delay thus showed a more rapid decrease with intensity than the recorded SFOAE data of Schairer et al. (2006).

Level-dependent compression characteristics such as the level-curve and the waveform fine-structure were adequately modeled, such that time-dependent properties of the CEOAE level-curve can be investigated with this model.

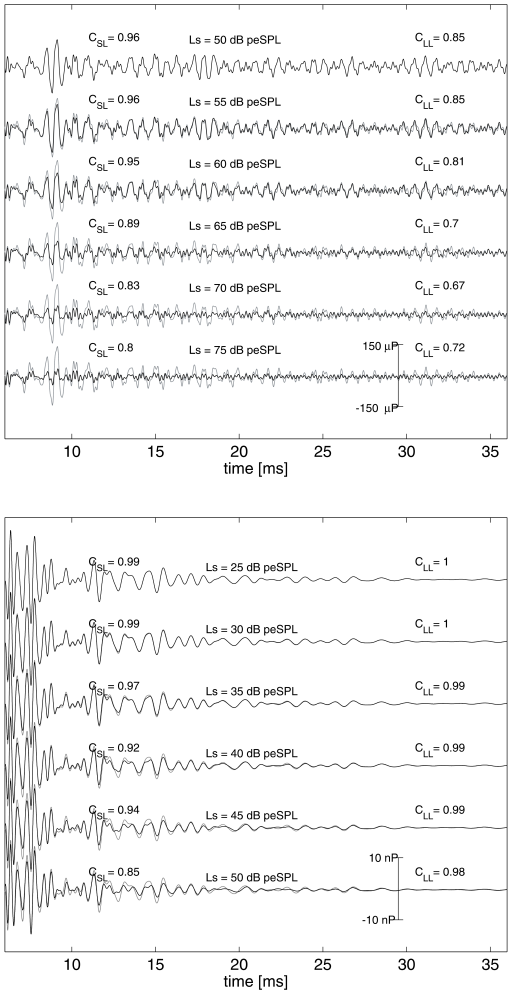


Figure 5.7: Stimulus-synchronization for the recorded (top) and simulated (bottom) CEOAE for an ICI of 2 ms. The US (gray traces) and DS response (black traces) from Fig. 3.2 were overlaid for fixed L_t and different L_s. L_t was 65 dB peSPL for the recorded CEOAE and 40 dB for the simulated CEOAE. The correlation factors were calculated for short-latency (6–18 ms; SL) and long-latency (24–36 ms; LL) windows of the CEOAE.

5.3.2 Stimulus-synchronization of the CEOAE

Chapter 2 addressed the need for stimulus-synchronization to occur when using temporal suppression as a measure for short-duration changes in cochlear compression. Stimulus-synchronization describes a "resetting" process where the phase of the recorded waveform is identical whenever a click is presented (i.e., in a measurement and at the same level). It is because of this property that one can filter the CEOAE from the background noise by averaging several hundreds of click recordings in a measurement (i.e., the CEOAE is correlated whereas the noise is not). If stimulus-synchronization was different for the test than for the suppressor click, then suppression as determined in Fig. 2.2 (method 2) and Fig. 3.2, could not be related to the CEOAE level-curve. The lack of stimulus-synchronization would introduce phase differences such that the derived suppressed (DS) response would show a different waveform pattern than the unsuppressed (US) response. Suppression, calculated as the level difference between the two responses, would then describe both a magnitude and phase effect. Whereas the magnitude component describes the change in compression caused by the ICI between the suppressor and test click, a phase component would be undesired as it biases the suppression measure. Stimulus-synchronization to the suppressor and test click was investigated in Fig. 5.7 for the recorded and simulated CEOAE. If the stimulus-synchronization mechanism would be intact when applying the measurement paradigm in Fig. 3.2, then a good correlation between the unsuppressed and derived suppressed CEOAE would be obtained. The correlation between the US and DS response of the recorded CEOAE was high for L_s below L_t, and decreased as L_s increased. This decrease in correlation was caused by the more coarse waveform of the DS response compared to the US response, for higher L_s. The simulated CEOAE showed good correlation factors for all L_s and L_t configurations tested. The model thus accounted well for the stimulus-synchronization that was observed in the recordings of subject #1. This implies that the suppression levels calculated for the model simulations were able to describe changes in compression.

5.3.3 Time scale of temporal suppression

Fig. 5.8 shows simulation results for experiment 1, where the time scale of temporal suppression was investigated for several L_t and L_s configurations. For every L_t, three L_s conditions were measured, and suppression was calculated for the short-latency (black traces) and long-latency (gray traces) region of the simulated CEOAE. Suppression was found to be maximal for coinciding clicks (i.e., ICI = 0 ms), with

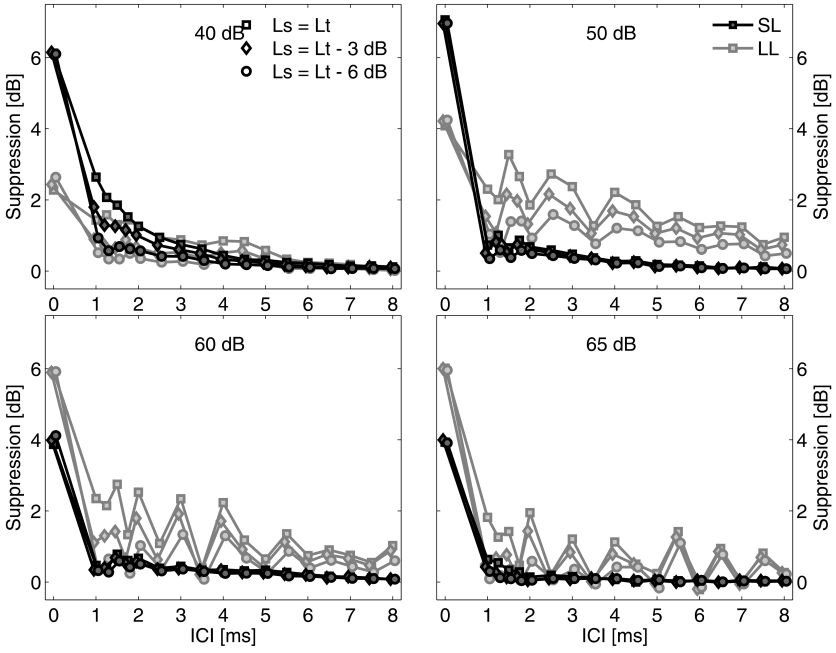


Figure 5.8: Simulated suppression levels [dB] for ICIs between 0 and 8 ms. Each panel represents a different L_t, and in every panel three L_s conditions were tested: L_s = L_t, L_s = L_t - 3 dB, and L_s = L_t - 6 dB. rms levels were calculated for the short- (6–18 ms; SL) and long-latency (24–36 ms; LL) time frames of the CEOAE.

values ranging between 4 and 7.5 dB for the short-latency CEOAE and between 2.5 and 6 dB for the long-latency CEOAE. For ICIs different from zero, suppression decreased with increasing ICI. Maximal suppression for non-coinciding clicks was found for an L_t of 40 dB peSPL in the short-latency CEOAE (2.5 dB at 1 ms), and

for an L_t of 50 dB peSPL in the long-latency CEOAE (3 dB at 1.5 ms). For all L_t conditions, suppression decreased with increasing L_s and the largest differences were observed for an L_t of 40 dB peSPL. ICIs between 0 and 1 ms were not tested in the model as temporal overlap in the transient stimuli (see Fig. 5.1) would for these ICIs bias the suppression measure. Stimulus overlap could occur until an ICI of 4 ms, but preliminary tests[1] have shown that for ICIs longer than 1 ms, this overlap did not influence the suppression measure. Suppression for the long-latency CEOAE showed overall higher levels than for the short-latency CEOAE, and the time scale of the effect was also longer for this latency region. A fluctuating suppression pattern was furthermore observed for the long-latency CEOAE when increasing the ICI.

There were two clear differences between the simulated suppression levels and the experimental data obtained for subject #1, both shown in Fig. 5.9. Firstly, suppression

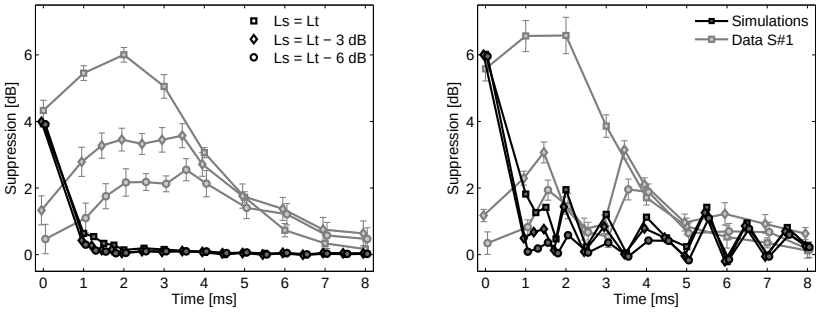


Figure 5.9: Suppression levels [dB] for subject #1 (gray curves) and model (black curves) and ICIs between 0 and 8 ms. L_t was 65 dB peSPL and three different L_s conditions were tested. rms levels were calculated for short- (left) and long-latency (right) time-frames of the CEOAE.

for the recorded CEOAE initially increased with ICI, whereas suppression decayed rapidly for the simulated results. The overall suppression levels were also more than twice as large for the experimental data than for the simulated results (i.e., for ICIs $\neq$ 0). Secondly, the fluctuations in suppression for increasing ICI, observed for the long-latency region of the simulated CEOAE, did not appear in the recordings of subject #1

[1] Experiment 1 was repeated for 80 μs Gaussian pulses. Suppression for this stimulus yielded quantitative similar results than in Fig. 5.8

(nor for the other subjects tested in chapter 3). It is clear from the comparison between data and simulations in Fig. 5.9, that the model did not account for the time scale of temporal suppression that showed maximal suppression at ICIs different from zero.

5.3.4 Temporal adaptation of the CEOAE level-curve

Unsuppressed and derived suppressed level-curves were compared for the recorded and simulated CEOAE in Fig. 5.10. Temporal adaptation, as the deviation from the

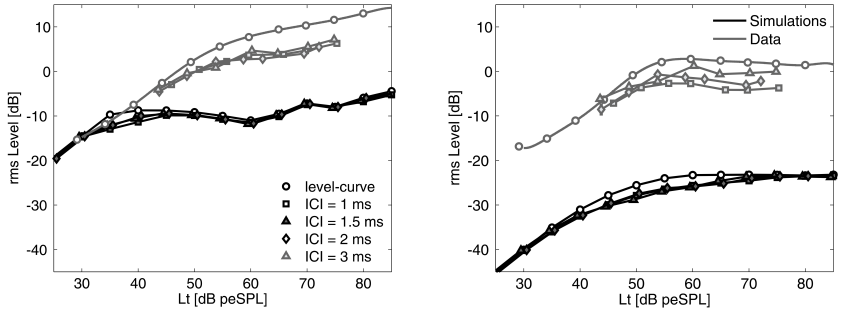


Figure 5.10: Simulated (black) and recorded (gray) for the short- (left) and long-latency (right) CEOAE level-curves. As the output pressure of the recorded and simulated CEOAE were not comparable, the simulated level-curves were vertically offset by 65 dB, such that the unsuppressed short-latency level-curve at 30 dB coincided with the experimental data. Suppressed level-curves were obtained for $L_s = L_t$, and for ICIs of 1, 2, 3 ms for the recorded CEOAE, and for ICIs of 1, 1.5 and 3 ms for the simulated CEOAE.

unsuppressed level-curve, was not observed in the same degree for the simulations than for the experimental data. For the short-latency results (left panel), the model simulations showed suppression that was concentrated near the knee-point of the unsuppressed level-curve at 35 dB peSPL. Suppression furthermore decreased to zero for L_t above 55 dB peSPL. The experimental results on the other hand, showed an offset of the suppressed short-latency level-curves for higher stimulation levels. The simulated long-latency results (right panel), that showed suppression up to simulation levels of 70 dB peSPL, agreed better with the experimental data. The suppressed level-curves for the recorded long-latency data were again seen as an offset from the unsuppressed level-curves for higher stimulation levels, whereas the simulated

unsuppressed and suppressed level-curves coincided for stimulation levels above 70 dB peSPL. The lack of suppression for stimulus levels above 70 dB peSPL was probably due to the linear behavior of the modeled CEOAE level-curve for these stimulus levels (see Fig. 5.5). Temporal adaptation was observed in the simulations near the knee-points of the level-curves, but neither the amounts of suppression nor the extent to higher intensities observed in the experimental data, were accounted for.

5.3.5 Discussion

Mechanisms underlying CEOAE generation

The peaks of the recorded CEOAE waveform showed both small latency leads and lags with increasing stimulus intensity. The simulated CEOAE on the other hand, showed a decrease in group delay (or a phase lead) in the waveforms for increasing intensities. Carvalho et al. (2003) showed recorded CEOAE waveforms similar to the ones presented here, and argued that the phase changes in the CEOAE could only occur if the origin of the CEOAE was found in inter-modulation distortion. They postulated that it was only possible to find intensity-dependent phase lags/leads in the CEOAE waveform if the CEOAE frequency components arose from locations other than the resonance frequency of the stimulus. Carvalho et al. (2003) furthermore stated that the intensity-dependence of CEOAEs did not comply with an intensity-invariance of the fine-structure in the BM impulse response, as recorded by Recio and Rhode (2000).

The model simulations presented in this study showed that intensity-dependent phase changes in the CEOAE were possible in a model where the fine-structure of local BM impulse responses was intensity-invariant[2]. Schairer et al. (2006) argued that the intensity-dependent SFOAE group delay decrease was linked to the leads observed at CF in BM phase transfer functions recorded to increasing stimulus intensities (Ruggero et al., 1997). The model simulations in Chapter 4 confirm this correlation

[2] The modeled BM impulse responses maintained a fine-structure invariance for intensities below 80 dB (see Fig. 4.9 in Chapter 4). The simulated CEOAEs showed a decrease in group delay for intensities above 55 dB (see Fig. 5.6)

between a latency decrease in the BM phase transfer function at CF (see right panel of Fig. 4.8), and a decrease in the CEOAE group delay for increasing stimulus intensities.

It is argued here that the intensity-dependence of the CEOAE waveform does not need to stem from inter-modulation distortion, but that it can be accounted for by a model based on coherent reflection. According to this idea, the resonance frequency (CF) of a specific location on the BM does not change with intensity, in agreement with results of Recio and Rhode (2000) and Shera (2001b). CEOAEs were modeled by placing fixed BM-irregularities along the length of the BM, ensuring that CEOAEs could only arise from cochlear locations where these irregularities were placed. It is postulated here that the global mechanical properties of the cochlea account for the intensity-dependent phase changes in the CEOAE, and not the local BM resonance frequency that remains invariant with intensity. As the stimulus intensity increases, a section of the BM can be dominated by driving frequencies that are lower than its CF, due to the (intensity-independent) frequency glide (Shera, 2001b) (see also left panel of Fig. 4.8 and Fig. 4.9). As a consequence, a half-octave shift in BF occurs in the local magnitude transfer functions when the stimulus intensity is increased from 0 to 90 dB. It is believed that this half-octave shift in BF, that goes together with a local phase-lead at CF in the local BM transfer functions, is the cause for both the intensity-dependent phase leads observed in the SFOAE group delay, and the phase leads in the CEOAE waveforms. Inter-modulation distortion is in this context not needed to account for the intensity-dependent phase leads in the CEOAE. It is however possible that there is a role for inter-modulation products in the exact phase-pattern of the CEOAE. The present model simulations and the experimental SFOAE data by Schairer et al. (2006) only showed a decrease in the group delay with increasing stimulus intensities. Conversely, Carvalho et al. (2003) showed both phase lags and leads in the CEOAE waveform of the same subject. It is possible that the existence of inter-modulation products would influence the otherwise decreasing group delay with increasing stimulus intensity.

Mechanisms underlying temporal adaptation in CEOAEs

Even though the model accounted well for intensity-dependent features of the CEOAE, it did not describe key features of temporal adaptation. The model simulations were

able to show temporal suppression, but as opposed to the experimental results, maximal suppression was found when the two clicks coincided and not when the ICI was 1–4 ms. Simulated suppression was furthermore shown to never exceed 3 dB for non-coinciding clicks, and was thus considerably smaller than for the experimental results that showed maximal suppression levels between 6 and 10 dB. Lastly, the sustained suppression observed in the suppressed level-curves for fixed ICIs and high test-click levels, was absent in the model simulations. It can be concluded that the transmission-line model, implemented with an instantaneous nonlinearity, did not account for important features of temporal adaptation in CEOAEs. This result leads to two hypothesis about the cochlear mechanisms that underly temporal adaptation.

Firstly, it is possible that temporal overlap of local BM impulse responses still provide a basis for temporal adaptation. However, an instantaneous nonlinearity that tracks the BM impulse response fine-structure cannot account for temporal adaptation. Rather than following the fine-structure of the y_{BM} or $\dot{y}_{BM}$ pattern, the nonlinearity could, for example, track the instantaneous *envelope* of excitation. For sinusoidal signals, such a nonlinearity could track the envelope function $y_{BM}^2 + \left(\frac{\dot{y}_{BM}}{CF}\right)^2$, and for click stimuli an instantaneous envelope function could look like $|y_{BM}| + \frac{|\dot{y}_{BM}|}{CF}$.

One reason to argue against an instantaneous nonlinearity that tracks the fine-structure of $\dot{y}_{BM}$, was found in the fluctuating suppression pattern observed for increased ICIs when analyzing long-latency suppression in the model. As the spectral content of the simulated long-latency CEOAE showed energy concentrated in the 1–1.5 kHz region, the fine-structure pattern in the BM impulse response could easily show a phase relation with the ICI. The fluctuations observed in suppression for increasing ICIs were most likely caused by the alternating constructive and deconstructive summing of the BM impulse response at the ICIs tested. Such a fine-structure in the suppression patterns was not observed in the experimental data, which supports the view that an instantaneous nonlinearity that tracks $\dot{y}_{BM}$ is not enough to account for temporal adaptation. One advantage of an envelope tracking nonlinearity over a nonlinearity that tracks $\dot{y}_{BM}$ is that the nonlinearity would sustain its compression factor, rather than turning on and off every time a zero-crossing in $\dot{y}_{BM}$ occurs. Even though the effect of this may be small when looking at single frequency, introducing

sustained compression across the whole BM, may lead to overall increased levels of compression that could potentially lead to higher suppression levels.

A second advantage of such an envelope tracking nonlinearity is that it may not be needed to introduce a time-dependent, or memory, factor in the nonlinearity function. As discussed earlier in chapter 3, the envelopes of the local y_{BM} impulse responses measured by Recio and Rhode (2000) reached their maxima at 1 to 2 ms after stimulus onset. Maximal overlap of two temporally separated local BM impulse responses would thus occur at ICIs different from zero, regardless of the nonlinearity strategy adopted. By using an instantaneous envelope tracking nonlinearity, the time-scale of temporal suppression could theoretically still be caused by the inherent delay in the maxima of the BM impulse response patterns that were observed in the experimental data of Recio and Rhode (2000), and in the simulated BM impulse response patterns shown in Fig. 4.9.

However, it is unclear how an envelope tracking nonlinearity alone can account for suppression maxima occurring at ICIs different from 0. Previous studies have attempted to model temporal suppression by simulating suppression with a symmetric 1 kHz tone-burst as the input to a local BM input/output function (Kapadia and Lutman, 2000a; Harte et al., 2005). Temporal overlap of local BM impulse responses (i.e., the symmetric tone-bursts) caused different compression properties at the output of the static nonlinearity in both models, during the time the symmetric tone-bursts overlapped. Both studies postulated that it was the increase in compression seen at the output of the local BM input/output function that would account for temporal suppression in CEOAEs. Whereas the model of Kapadia and Lutman (2000a) used a fine-structure tracking nonlinearity, the model of Harte et al. (2005) used an envelope tracking nonlinearity. In both of the proposed models, maximal suppression was obtained for coinciding clicks when static nonlinearities were adopted. The difference between the models of Kapadia and Lutman (2000a), Harte et al. (2005) and the simulations presented here, is that they did not account for broadband cochlear effects, nor for asymmetric BM impulse responses. Due to the cochlear impedance function, the BM impulse responses in the present study showed asymmetric envelopes with maxima near 1–2 ms.

Secondly, it is possible that another mechanism, such as an adaptation process

in the outer hair cells, is responsible for temporal adaptation in CEOAEs. Fast adaptation in outer hair cells is manifested as a saturation of the gating channel (i.e., the K^+/Ca^{2+} transduction channel in the cilia of the hair cell) that occurs when deflecting the hair cell from its resting state (Hudspeth and Gillespie, 1994). This adaptation process has been demonstrated in vitro in isolated outer hair cells, and was shown to last about 10–20 ms (Hudspeth and Gillespie, 1994). However, to date there is no evidence that this adaptation process also takes place in the human cochlea. If all OHCs along the length of the cochlea would, when excited by a click, show a 10 ms adaptation period before returning to their resting state, then this may influence the compression factors within this time period. If during the first 10 ms after click excitation, a new click is presented, the outer hair cells could then show different compression factors depending on the ICI and the saturation state they are in. A fast adaptation process can be implemented in the model by including a memory into the nonlinearity. This memory function would comprise a feedback mechanism such that compression could be sustained for 10 ms after stimulus onset. Introducing a delay function in the local compression factors may cause temporal suppression to show maxima at ICIs different from zero. Harte et al. (2005) proposed a single channel model based on a local BM input/output function, that introduced a delay before the gain was applied to the input. As a consequence of the delayed compression, maximal suppression for symmetrical tone-bursts was obtained for ICIs different from zero. It is however unclear where such a delay in the applied compression would physiologically originate from. The memory function that is proposed in this study would dynamically shift the BM I/O function horizontally within the first 10 ms after stimulus application. The local BM I/O functions would then show similar behavior to the OHC input/output functions described in Hudspeth and Gillespie (1994).

It is possible that temporal adaptation in CEOAEs supports the existence of a short-term adaptation process in human hearing, but more evidence is needed to support this hypothesis. It remains to be investigated how these adaptation processes would influence local BM compression properties, and if they are really due to a saturation process in the OHC cilia.

5.4 Conclusion

The presented model, implemented with an instantaneous nonlinearity that tracked local $\dot{y}_{BM}$ impulse responses, accounted well for intensity-dependent features of the CEOAE. Temporal adaptation was furthermore modeled well for the linear and knee-point region of the CEOAE level-curve, where both the recorded and simulated results showed increased suppression for increased stimulus levels. The model simulations did, however, not account for the sustained suppression levels that were observed in the compressive region of the experimental CEOAE level-curves, nor for the time scale of the effect (i.e., maximal suppression at ICIs different from 0). The model, implemented with an instantaneous nonlinearity that tracked $\dot{y}_{BM}$ was thus inadequate to model important features of temporal adaptation of the CEOAE level-curve. Temporal overlap of BM impulse responses could still be the basis of temporal adaptation, if for example, the nonlinearity would track the instantaneous envelope of the impulse responses rather than their fine-structure pattern. Alternatively, another mechanism such as outer hair cell adaptation may underly temporal adaptation. OHC adaptation would be implemented as a memory or feedback process in the local BM input/output functions. By ruling out an instantaneous y_{BM} and $\dot{y}_{BM}$ tracking nonlinearity as the mechanism underlying temporal adaptation in CEOAEs, this study has ruled out one of the possible hypotheses on the underlying cochlear mechanisms.

6

Overall Summary and Discussion

In this thesis, recorded and simulated click-evoked otoacoustic emissions were analyzed to investigate dynamic features of the nonlinear processing taking place in the human cochlea. The main objective was to investigate whether short-time (0–20 ms) dynamic features in cochlear nonlinearities, that have been found to exist in single OHC recordings in animals (see Martin (2007) for a review), are also present in the intact human system. As the nonlinear properties of CEOAEs arise from nonlinearities present in OHC behavior (see Kemp (2007) for a review), CEOAEs were used as a non-invasive technique to investigate this. In the 80's, Kemp and Chum (1980) observed that the CEOAE amplitude reduces when a second (i.e., suppressor) click is presented close in time to the evoking click. This experiment was the first to suggest a short time-constant (0–10 ms) related to the generator mechanisms underlying the OAE, and the original experiment has been repeated and extended several times after that (Tavartkiladze et al., 1994; Lina-Granade and Collet, 1995; Kevanishvili et al., 1996; Kapadia and Lutman, 2000b; Hine and Thornton, 2002). As it remained unclear how dynamic changes in cochlear nonlinearities can account for the reported amplitude alterations in the CEOAE, this thesis investigated the cochlear mechanisms underlying the results.

Chapter 2 addressed the experimental challenges that arise when using CEOAEs to investigate dynamic features of cochlear compression. When two clicks are presented close in time to each other, a CEOAE to the combined stimulus is recorded because the CEOAE to the first click (i.e., suppressor) overlaps in time with the CEOAE to the second (i.e., test) click. The CEOAE to the test click thus needs to be "derived" from the response recorded to the double click stimulus to investigate amplitude reductions of the CEOAE due to the temporal spacing between the clicks. Historically, several techniques have been adopted to derive this response, leading

to somewhat conflicting results. In Chapter 2, it was investigated to what extent the measurement methods themselves were responsible for the discrepancies in the reported data, and it was found that the suppression measure consisted of a phase and magnitude component. The magnitude component was caused by a change in the underlying compression mechanisms, and the phase component originated from waveform changes of the CEOAE. A phase component in the suppression measure is undesired, as suppression should describe magnitude changes in the CEOAE caused by compression changes in the underlying mechanisms. It was demonstrated that an undesired phase component was present in the measurement methods adopted by Tavartkiladze et al. (1994), Kevanishvili et al. (1996) and Hine and Thornton (2002). The augmentation (i.e., negative suppression) they reported was thus probably due to this phase component, rather than being caused by a gain increase in the underlying cochlear mechanisms. An undesired phase component was furthermore found to occur when stimulus-synchronization between the CEOAE and the evoking click was absent. Stimulus-synchronization describes a "resetting" process where the phase of the recorded CEOAE waveform is identical, whenever a click is presented (i.e., in a measurement and at the same level). If stimulus-synchronization is different for the test than for the suppressor click, then waveform changes can occur in the derived CEOAE to the test click, leading to a phase component in the suppression measure. The experimental results in Chapter 2 and 3 demonstrated that stimulus-synchronization was intact when the suppressor was presented at equal or lower levels than the test click. For these stimulus levels, and with the measurement paradigm introduced by Kemp and Chum (1980), suppression in CEOAEs provided a measure for investigating dynamic changes in cochlear nonlinearities.

In Chapter 3, an optimized recording technique was adopted to investigate dynamic changes in the CEOAE level-curve (i.e., the relation between stimulus and CEOAE level), when presenting a suppressor close in time before a test click. It was hypothesized that, if the CEOAE level-curve to the test and suppressor click were identical when presented alone, the level-curve to the test click would change when the suppressor was presented 0–10 ms before the test click. This change in the CEOAE level-curve was named "temporal adaptation" of the CEOAE level-curve and was demonstrated for all four subjects tested in Chapter 3. Temporal adaptation

was shown to have different features for the linear and compressive region of the CEOAE level-curve. Near the knee-point of the level-curve, suppression increased with increasing stimulus level for a fixed inter-click interval (ICI). In the compressive region of the CEOAE level-curve, temporal adaptation was observed as a vertical offset from the level-curve towards lower CEOAE levels. Temporal adaptation was maximal for ICIs between 1 and 4 ms, depending on the subject, and disappeared for ICIs larger than 8 ms. Maximum suppression levels were obtained in the compressive region of the CEOAE level-curve, and ranged between 6 and 10 dB for the different subjects. Temporal adaptation was furthermore shown to be similar for the purely evoked (i.e., short-latency; 6–18 ms) emission component, and the SSOAE (i.e., long-latency; 24–36 ms) component in the CEOAE.

The experimental results in Chapter 3 attributed temporal adaptation of the CEOAE level-curve to changes in local place-fixed BM mechanisms, and the reasoning for this statement was twofold. Firstly, temporal adaptation was shown to affect the short- and long-latency CEOAE level-curve similarly. At the same time, the level-curves showed greater amounts of compression for the long-latency region associated with SOAEs, than for the short-latency region. Thus, even though the difference between the generator mechanisms of evoked and SOAE emission components was reflected in the CEOAE level-curve, no substantial differences were observed in how temporal adaptation affected these level-curves. Temporal adaptation should thus be linked to an underlying mechanism that is common to both types of emissions. Following the taxonomy of Shera and Guinan (1999), CEOAEs arise as coherent reflections of the forward traveling wave to local place-fixed BM irregularities. SOAEs were in this taxonomy explained as amplitude stabilized cochlear standing waves that arise through narrowband coherent reflection taking place even when no external stimulus is applied (Zweig and Shera, 1995; Shera and Guinan, 1999). If the origin of temporal adaptation is to be found in a mechanism that is shared between both types of emissions, then the local place-fixed BM mechanisms giving rise to coherent reflection, emerge as logical candidates. The second argument that links temporal adaptation to local place-fixed BM mechanisms, was found in the fluctuation of temporal adaptation across the CEOAE waveform. Shera and Cooper (2009) investigated the origins of the "waxing" and "waning" periods of the CEOAE waveform, i.e., the multiple temporal lobes in the

CEOAE waveform. They argued that the waxing periods were associated with place-fixed BM irregularities, and that the waning periods were due to reflections of energy from the stapes. It was demonstrated for one subject in this study, that suppression was larger for the waxing periods than for the waning periods of the CEOAE. This finding supports the origin of temporal adaptation to lie in local place-fixed BM mechanisms.

The experimental data presented in Chapter 3 did not lead to a conclusion about the exact nature of the local place-fixed BM mechanisms underlying temporal adaptation of the CEOAE level-curve. Historically, two hypotheses have been proposed on the local BM nonlinearities underlying temporal adaptation in CEOAEs. In the first, the nonlinearity was assumed to be static or time invariant (Kemp and Chum, 1980; Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Hine and Thornton, 2002; Harte et al., 2005). Temporal adaptation would in this framework be caused by a constructive summing of local BM impulse response patterns. The duration of these impulse response patterns (Recio and Rhode, 2000) causes them to overlap in time when two clicks are presented closely in time to each other. If constructive summing takes place within this overlap period, then the input to the nonlinearity would be larger than for a single click. As a result, more compression would be applied by the static nonlinearity, which would lead to increased suppression levels. In the second hypothesis, a time dependence that operates on a time scale of 0–10 ms would exist in the local BM nonlinearities (Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Verhulst et al., 2008). In this framework, the BM input/output function would dynamically shift horizontally for a duration of 0–10 ms after a click is applied. When a second click is presented within this time period, it would operate from a different BM input/output function than the first click, whereby different compression characteristics would be obtained. The idea of introducing a dynamic process in the local BM nonlinearities was based on OHC adaptation that takes place in isolated OHCs on a time scale of 0–20 ms (Hudspeth and Gillespie, 1994). The experimental data in Chapter 3 were inconclusive as to which of the two local BM mechanisms (or a combination of both) would underly temporal adaptation in CEOAEs. For the compressive region of the CEOAE level-curve, both strategies could lead to the same suppression levels. For the linear region and near the knee-point of the CEOAE level-curve, the adaptive nonlinearity seemed less likely to underly the results. Suppression

for fixed ICIs was shown to increase with stimulus level in the linear region and near the knee-point of the CEOAE level-curve. The knee-points of the "unsuppressed" and "adapted" level-curves were furthermore found to occur for the same stimulus level. If an adaptive nonlinearity would underly the suppression results, then it would be more likely that the whole CEOAE level-curve shifted horizontally with ICI, also for the linear region.

The presented CEOAE data led to a comprehensive description of dynamic features of the CEOAE level-curve. However, to find the origin of the effect, a link between dynamic changes in local BM input/output functions and dynamic features of the broadband CEOAE level-curve needed to be established. One way of doing this, would be to record animal BM impulse responses to closely spaced clicks. In an alternative non-invasive approach, the CEOAE results could be simulated with a model of the cochlea that accounts for broadband processing.

In this thesis, a time-domain model of the cochlea that produces CEOAEs was adopted to investigate the origin of the temporal adaptation. The model, presented in Chapter 4, was developed to account for features observed in BM impulse responses recorded in chinchilla (Recio and Rhode, 2000). A first feature comprised the increased duration of the BM impulse response when going from base to apex in the cochlea. Secondly, the model accounted for the observed intensity-invariance of the zero-crossings of local BM impulse responses. The model was used to test the first hypothesis on the mechanisms underlying temporal adaptation in CEOAEs (i.e., temporal overlap of BM impulse responses). This was achieved by ensuring that the simulated BM impulse response behavior was in agreement with the experimental data of Recio and Rhode (2000). The implemented nonlinearity was furthermore chosen to track the BM impulse response fine-structure instantaneously (i.e., time-invariant nonlinearity), whereby the amount of compression applied in one time step depended on the BM velocity of the previous time step.

The results in Chapter 5 demonstrated that the developed model was able to simulate level-dependent features of the CEOAE well. Firstly, a reduction in group delay for increasing intensities was obtained, in agreement with the recorded SFOAE data by Schairer et al. (2006). Secondly, the recorded and simulated CEOAE level-curve showed similar compressive behavior for the short-latency region of the CEOAE. As

the reflectivity of the middle-ear impedance of the model was kept low, the model did not account for strong SOAEs in the long-latency region of the CEOAE. The model was able to simulate some degree of temporal adaptation, but did not account for several key features observed in the experimental data: maximal suppression in the model was found when the suppressor and test click coincided (ICI = 0 ms), whereas the experimental data showed suppression maxima for ICIs between 1 and 4 ms. Suppression decreased drastically for ICIs larger than 0 ms and never exceeded 3 dB when the suppressor and test clicks had equal levels. The experimental results showed larger suppression levels with maxima in the range of 6 to 10 dB. Temporal adaptation was modeled well for the linear and knee-point region of the CEOAE level-curve, where both the recorded and simulated results showed increased suppression for increased stimulus levels. The model simulations did, however, not account for the sustained suppression levels that were observed in the compressive region of the experimental CEOAE level-curves.

The presented model, implemented with a time-invariant nonlinearity that tracked the BM impulse response fine-structure, was thus unable to account for important features of temporal adaptation. These results led to two hypothesis about the cochlear mechanisms underlying the effect. Firstly, an instantaneous nonlinearity could still provide a basis for temporal adaptation. In this strategy, the nonlinearity would not track the fine-structure of the BM impulse response, but rather its envelope. Tracking the BM impulse response fine-structure leads to applied compression that fluctuates with the waveform. A more sustained compression across the duration of the BM impulse response would be obtained when tracking the envelope of the impulse response. The difference between a nonlinearity tracking the fine-structure or the envelope of the BM impulse response may not make a substantial difference when comparing suppression in narrow frequency bands (few sections on the BM), but it may make a difference for the broadband cochlear response. There is a possibility that the constructive and deconstructive summing of BM impulse responses, caused by the relation between the resonance frequency of the section and the ICI, may have canceled out the suppression effect along the length of the BM when tracking the BM impulse response fine-structure. It remains to be tested whether an envelope-tracking nonlinearity would perform better in this situation, and whether such a nonlinearity

could account for the relatively high levels of suppression (6 to 10 dB) found in the experimental data.

In a second hypothesis about the generator mechanisms underlying temporal adaptation, the nonlinearity itself would show a time dependence. This time dependence would look like a dynamic shift of the operating point on the BM input/output function in the time range of 0–10 ms. Alternatively, it can be seen as a horizontal shift of the BM input/output function in the first 0–10 ms after a click is applied. This idea stems from OHC adaptation, a time dependence present in isolated OHC (Hudspeth and Gillespie, 1994), and was earlier proposed as a possible mechanism underlying temporal adaptation (Lina-Granade and Collet, 1995; Kapadia and Lutman, 2000b; Verhulst et al., 2008). A dynamic shift in local BM input/output functions in the time period of 0–10 ms would be able to account for high suppression levels, and suppression maxima occurring at ICIs different from zero. There is, however, no physiological basis for implementing such dynamic processes in the gain mechanisms of the presented model. It is also unclear to what extent the local BM input/output function would shift depending on ICI, and whether the time scale of the effect would be similar along the whole length of the cochlea.

It can thus be concluded that the exact origin of temporal adaptation in the CEOAE level-curve has not been found yet. This thesis has contributed by describing temporal adaptation of the CEOAE level-curve, also at lower stimulus levels. In addition, several caveats in the measurement paradigms that have historically been used to investigate temporal adaptation were discovered. The approach of using a time-domain model of the cochlea has led to the exclusion of one of the possible hypotheses on the origins of temporal adaptation. Continuous improvement of the model may in the future help determine the origin of temporal adaptation in CEOAEs. Animal BM impulse response recordings to clicks presented close in time could help investigate cochlear dynamics involved in temporal adaptation.

A

Investigating Dynamic Changes in Cochlear Compression

A.1 Investigating dynamic changes in cochlear compression

It is important that the measurement paradigm in Fig. 3.2 results in suppression values representative of temporal changes in compression caused by cochlear mechanisms. A series of small control experiments was conducted to optimize the design of the experiment and to quantify different biases that can arise when using suppression in CEOAEs as a measure for cochlear compression changes. The potential biases on suppression were related to stimulus presentation and measurement implementation. The control experiments were conducted both in the ear canal of subject #1 and in a BK-2012 ear canal coupler, attached to a BK-4157 artificial ear. Clicks were presented at 65 dB peSPL and the ICIs tested were: 0.2, 0.33, 0.5, 1, 2, 3.33, 5, 6, 7, 8 ms.

Stimulus presentation

The effective stimulus level at the ear drum can be different than the intended stimulus level due to three reasons: (i) transducer hysteresis, (ii) temporal overlap of the suppressor and test click responses in the ear-canal, and (iii) real-ear-to-coupler differences. It is important that the effective stimulus levels at the ear drum are close to the intended levels, as the rarefaction and condensation click in DC and DCI are subtracted to form the DS response (see Fig. 3.2). Any effective stimulus level difference between DC and DCI would be observed as an artifact in DS at the suppressor location in DC and DCI. Alternatively, this artifact could be observed

as an OAE component in DS that depended on the difference between the effective absolute levels of L_t in DC and DCI. The OAE component would arise because the different stimulus levels in DC and DCI would evoke OAEs from different places on the CEOAE level-curve. Subtracting these CEOAEs would lead to an artifactual residue that can easily be confused with an actual OAE.

Transducer hysteresis in the ER-2 loudspeakers is important when playing condensation clicks followed by rarefaction clicks. The hysteresis in the ER-2's is such that 200 ms are required after a the presentation of condensation click to generate an equal level rarefaction click (see Kapadia et al. (2005)). The influence of transducer hysteresis on the effective stimulus levels was investigated by presenting the stimuli over either one, or two, ER-2 loudspeaker channels. When the stimuli were presented over two ER-2 channels, it was ascertained that condensation clicks were sent over one channel and rarefaction clicks over the other. Effective click levels were calculated for different ICIs as the peak-to-peak level of the recorded suppressor and test clicks. A maximum difference of 0.2 dB peSPL was found between the recorded click levels for the single versus double ER-2 channel recordings, for all 10 ICIs tested. Thus, transducer hysteresis would unlikely lead to a substantial artifactual OAE component. This was confirmed by the suppression levels calculated for subject #1 in Table A.1 that showed that suppression was not significantly different for the two loudspeaker channel configurations. Even though there was only a small deviation in effective

Table A.1: Suppression levels (S) in [dB] for the short-latency CEOAE (6–18 ms) of subject #1 in two different loudspeaker configurations. In the single channel configuration all clicks were sent over one ER-2 channel, in the double channel configuration the condensation clicks were sent over the first channel and the rarefaction clicks over the second channel. L_s and L_t were 65 dB peSPL, and the ICI tested were between 0 and 8 ms.

ICI [ms]	0.2	0.33	0.5	1	2	3.33	5	6	7	8
S single [dB]	4.4	3.3	4.5	5.6	6.5	4.3	1.6	1	0.4	0.1
std	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5
S double [dB]	4.3	2.7	4.5	5.6	6.1	4.4	1.1	0.3	-0.1	-0.3
std	0.6	0.5	0.5	0.6	0.6	0.5	0.5	0.5	0.5	0.5

stimulus levels that can be attributed to transducer hysteresis when presenting condensation and rarefaction clicks over one loudspeaker channel, it was decided to present the condensation and rarefaction clicks over two loudspeaker channels in this study.

The minimal influence of transducer hysteresis on the suppression results in Table A.1 indicates that suppression for preceding suppressors can be compared to studies that used a single loudspeaker channel for stimulus presentation (e.g. Kapadia and Lutman, 2000b; Verhulst et al., 2008).

The influence of ear-canal impedance on the stimuli was investigated using 80-μs clicks that were presented to the BK-2012 ear canal coupler, attached to a BK-4157 artificial ear. The levels of the recorded clicks (i.e., impulse responses) showed that there was no influence of the coupler (i.e., standard ear canal) on the effective click levels of the stimuli for ICIs longer than 1 ms. For ICIs shorter than 1 ms, the click responses overlapped in time in the coupler, giving rise to effective stimulus levels that were up to 2 dB different from the intended 65 dB peSPL. This L_t difference in the DC and DCI stimuli could give rise to an artifact in the DS response, and this was investigated for subject #1. It was found that for ICIs of 0.2 ms and 0.5 ms, the recorded test-click levels were deviating by -0.5 and -2 dB for the DC condition and by +1.5 and +1.5 dB for the DCI condition respectively. The non-monotonic suppression increase for ICIs between 0.2 and 1 ms observed in Table A.1 could be due to this suppressor stimulus mismatch. The observation of a similar "suppression dip" present for ICIs between 0.2 and 1 ms for all the subjects used in Verhulst et al. (2008) supports the existence of an artifact. Suppression for ICIs between 0 and 1 ms thus consists of a component caused by an effective stimulus-level difference, as well as a component caused by nonlinear interaction in the cochlea (i.e., the desired response). ICIs between 0 and 1 ms were not tested in this study to circumvent this problem.

The click stimuli used in this study were calibrated to the intended stimulus level with a BK-2012 ear canal coupler. The ER-10B measurement probe was placed with a tight fit into this cavity, which made that the probe tip was placed directly against the BK-4157 artificial ear calibrator. The effective stimulus level presented at the ear drum of a subject is not always identical to the calibrated stimulus level due to the quality of the probe fit. Even if there is no leakage, the measured stimulus level can deviate from the intended stimulus level, because of the small air gap between the probe tip and the ear drum (Scheperle et al., 2008). Even though the air gap between probe and ear drum was minimized, small effective level deviations across different

measurement sessions could not be avoided. The probe fit (with a sealed ear canal) did not affect the suppression measure because suppression investigates the *change* of an emission for similar probe fit conditions, but it might have influenced the effective stimulus levels presented. To compensate for across measurement stimulus deviations, an in-ear-calibration technique or a stimulus-level compensation can be adopted. This study accounted for effective stimulus-level differences by referencing all suppression results to the levels of the recorded clicks, rather than to the calibrated stimulus level. This level compensation ensured that the measurements were independent of the probe-fit of a specific session, and that suppression results obtained from different sessions could be compared to each other.

Implementation of the measurement paradigm

The stability of the reference US recording during a measurement session is crucial for the implementation of the measurement paradigm in Fig. 3.2. The frequency content of CEOAEs is known to be relatively stable across time as long as no hearing damage occurs, but the amplitude of CEOAEs is prone to fluctuations in the measurement conditions. Examples are the use of ototoxic drugs, recent noise exposure, middle ear pressure changes and more importantly for this study, probe fitting differences (Probst et al., 1991; Scheperle et al., 2008). As long as these measurement conditions do not considerably change over the course of a one-hour session, it is assumed that the CEOAE amplitude does not change, making it acceptable to use a single reference US recording for all ICI conditions measured. The stability of the US reference during a measurement session was investigated for two measurement implementations in Table A.2 for subject #1. Suppression for ref.1 was calculated for all ICIs conditions with an US reference that was recorded at the start of the measurement session. Correspondingly, suppression in the ref.2 condition was referenced to an US recording at the end of the measurement session. Furthermore, DC was recorded and averaged, then DCI, for every ICI condition tested in the ref. condition. The second implementation (int. in Table A.2) ensured the same measurement conditions for US, DC and DCI by interleaving the individual US, DC and DCI epochs. In the post-processing stage, the DCI averages were subtracted from the corresponding DC averages before the derived DS was averaged. In this second implementation, suppression was calculated with

Table A.2: Suppression levels (S) in [dB] for the short-latency (SL) and long-latency CEOAE (LL) of subject #1 for different implementations of the measurement paradigm. For ref.1, US was measured and averaged at the start of the measurement; for ref.2, US was measured and averaged at the end of the measurement. For every ICI, DC was recorded and averaged, where after DCI was recorded and averaged. For int, US, DC and DCI were presented interleaved such that for every ICI, an US, DC, DCI response was recorded before the next epoch was presented and averaged. L_s and L_t were 65 dB peSPL, and the ICI tested were between 0 and 8 ms.

ICI [ms]	0.2	0.33	0.5	1	2	3.33	5	6	7	8
S (SL) ref.1 [dB]	4.3	2.7	4.5	5.5	6.1	4.4	1.1	0.3	-0.1	-0.4
S (SL) ref.2 [dB]	5.7	4.1	6.0	7.0	7.5	5.8	2.5	1.7	1.3	1.0
S (SL) int [dB]	4.4	3.2	4.6	5.5	6.0	4.2	1.7	0.7	0.3	0.2
std (SL)	0.6	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5
S (LL) ref.1 [dB]	1.9	-1.7	2.4	1.6	1.6	0.3	-3.5	-3.8	-4.3	-4.5
S (LL) ref.2 [dB]	8.3	4.7	8.9	8.1	8.1	6.7	2.9	2.6	2.2	2.0
S (LL) int [dB]	6.3	3.6	7.0	5.9	7.7	3.2	0.7	0.2	-0.2	-0.4
std (LL)	0.2	0.3	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.3

averaged DS and US responses, measured at the same time for one ICI condition. Thus, any biological changes in experimental condition that could have occurred in DC and DCI, would also have occurred in US.

Suppression obtained for the ref.1 and ref.2 conditions in Table A.2 had similar increasing levels up to an ICI of 2 ms, where after suppression decreased for larger ICIs, but the suppression patterns were offset from each other. Suppression calculated with ref.2 showed levels that were about 1 dB higher for the short-latency CEOAE and about 5 dB higher for the long-latency CEOAE. A visual inspection of the waveforms of the averaged US in ref.1 and ref.2 (not shown), indicated that the amplitude of the waveform increased from start to end of the experiment. Broadband frequency spectra of these US recordings showed furthermore that the spectral content remained the same. The amplitude deviation of the US waveforms in ref.1 and ref.2 was observed to be larger for the long-latency CEOAE than for the short-latency CEOAE. This was in agreement with the offset observed between the ref.1 and ref.2 condition for the long-latency data of Table A.2. The correspondence between the US amplitude variation and suppression offset, led to the conclusion that suppression was biased when only using one US reference. The interleaved recording technique minimized any bias caused by varying measurement conditions. Furthermore, even if SOAE activity would vary in strength across the measurement, an interleaved recording

technique would still be able to describe suppression as the relative change between the SSOAE in the unsuppressed and suppressed condition. Suppression for the int. implementation in Table A.2 were situated in between the values reported for the ref.1 and ref.2 conditions. In light of this preliminary study, all ICI and level configurations were recorded using an interleaved recording technique.

Kapadia and Lutman (2000a,b) adopted a semi-interleaved technique, where 500 averages were recorded for the unsuppressed (US) and suppressed (DC and DCI) condition for every ICI and level configuration. A corresponding US condition was measured and averaged after every averaged DC-DCI condition, which means that a small bias due to varying measurement conditions could have occurred. Kapadia and Lutman (2000a,b) showed mainly average suppression data and short-latency results, which makes the extent of this potential bias on the presented results would be minimal. The suppression results presented in Verhulst et al. (2008) used a single US reference for all ICI conditions in a session, thus suppression levels could have been contaminated by varying measurement conditions. Even though the existence of such biases on calculated suppression levels is hard to prove, the results presented in Table A.2 show that they can cause significant offsets of the suppression measure. Therefore, these effects have to be taken into account when interpreting suppression data in terms of cochlear compression changes. Verhulst et al. (2008) reported augmentation (i.e. negative suppression levels) for the long-latency CEOAE of one subject in their study. Since the long-latency CEOAE is prone to biases caused by using a single reference US for all tested ICI conditions, these results should be interpreted cautiously. Neither Kemp and Chum (1980) nor Kapadia and Lutman (2000a,b) reported augmentation, and it is thus likely that the augmentation reported for the long-latency CEOAE for one subject in Verhulst et al. (2008) was due to a bias caused by varying measurement conditions, rather than to decreased compression in the underlying cochlear mechanisms.

B

Derivations of the Transmission-Line Model Equations

B.1 The traveling wave equation

The BM acceleration $\ddot{y}$ and the pressure balance at every section n of the model in Fig.4.1 were derived.

B.1.1 The pressure p_n

First, an expression was found for the pressure p_n acting on section n of the cochlea. From the general transmission-line expression of the shunt impedance Z_p in equation 4.4, and the specific expression of $Z_p(s)$ in equation 4.6, it follows that:

$$Z_p(s) = p_n \frac{dx}{U_n} = \omega_{c0} M_{p0} \left(s^2 + \delta s + 1 + \rho \mathrm{e}^{-\psi s}\right)/s\,. \tag{B.1}$$

The volume velocity U can be expressed as:

$$U = b\,dx\,\dot{y}\,, \tag{B.2}$$

where b represents the width of the basilar membrane and where dx represents the length of a section of the basilar membrane. From equation B.2, an expression for $\frac{dU}{dx}$ is found as:

$$\frac{dU}{dx} = b\,\dot{y}\,. \tag{B.3}$$

After replacing the scaling parameter s by $j\omega/\omega_{cn}$, and after applying equation B.3 to equation B.1, an expression for the pressure p_n can be written as:

$$p_n = b\,\omega_{c0} M_{p0} \left[\frac{1}{\omega_{cn}} \ddot{y}_n + \delta \dot{y}_n + \omega_{cn} \left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{cn}}} \right) y_n \right] . \tag{B.4}$$

B.1.2 The basilar-membrane acceleration $\ddot{y}_n$

The basilar-membrane acceleration $\ddot{y}_n$ for a single section n can be written in function of the pressure p_n, the basilar-membrane velocity $\dot{y}_n$ and displacement y_n of that section. This is done by rewriting equation B.4 to $\ddot{y}_n$.

$$\ddot{y}_n = \frac{\omega_{cn}}{b\,\omega_{c0} M_{p0}}\, p_n - \omega_{cn}\delta \dot{y}_n - \omega_{cn}^2 \left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{cn}}} \right) y_n \tag{B.5}$$

This equation can be simplified by introducing a factor q_n and g_n, such that equation B.5 becomes:

$$\ddot{y}_n = q_n - g_n \,, \tag{B.6}$$

where:

$$q_n = \frac{\omega_{cn}}{b\,\omega_{c0} M_{p0}}\, p_n \text{ with } [q_n] = \left[\frac{L}{T^2} \right] , \tag{B.7}$$

and

$$g_n = \omega_{cn}\delta \dot{y}_n + \omega_{cn}^2 \left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{c(n)}}} \right) y_n \text{ with } [g_n] = \left[\frac{L}{T^2} \right] . \tag{B.8}$$

The damping and stiffness can be defined from equation B.8 as the factors that are multiplied with the basilar-membrane velocity and displacement respectively:

$$D = \omega_{cn}\delta \text{ with } [D] = \left[\frac{1}{T} \right] \tag{B.9}$$

$$S = \omega_{cn}^2 \left(1 + \rho \mathrm{e}^{\left(\frac{-j\omega\psi}{\omega_{cn}} \right)} \right) \text{ with } [S] = \left[\frac{1}{T^2} \right] \tag{B.10}$$

Note here that both D and S vary along the length of the cochlea. They both decrease towards the apex of the cochlea if the damping constant $\delta > 1$ (i.e., passive model implementation). If $\delta < 1$, D increases exponentially towards zero from the base to the apex (i.e., active model implementation).

B.1.3 The pressure balance at section n

The pressure balance at a specific section n along the length of the cochlea was derived by applying Newton's second law in the horizontal direction of the schematic in Fig.4.1. This led to the series impedance Z_s that was given in equation 4.5. Before applying equation 4.5 to the schematic given in Fig. 4.1, the series impedance was rewritten with an expression of M_{sn} rather than M_{s0}. The relation between these two masses was given by Zweig (1991) as:

$$M_{sn} = \frac{\omega_{c0}}{\omega_{cn}\, M_{s0}}\,. \tag{B.11}$$

This means that the series impedance in equation 4.5 can be rewritten into:

$$Z_s = -j\omega M_{sn}\,. \tag{B.12}$$

With use of equation 4.3, equation B.12 can be rewritten to show a relation between the pressure difference dp over a specific section of the basilar membrane and the series mass M_{sn} corresponding to that section.

$$dp_n = -dx\, M_{sn}\, \dot{U}_n \tag{B.13}$$

Applying equation B.13 to the schematic in Fig.4.1 leads to the following two expressions:

$$p_n - p_{n-1} = -dx\, M_{sn}\, \dot{U}^+_{n-1}\,, \tag{B.14}$$

$$p_{n+1} - p_n = -dx\, M_{sn}\, \dot{U}^+_n\,. \tag{B.15}$$

Equation B.14 and B.15 can be combined to write the pressure balance over a specific section n of the basilar membrane. This is done by applying the conservation of mass principle on node n of the schematic in Fig.4.1:

$$\dot{U}_n = \dot{U}^+_{n-1} - \dot{U}^+_n\,. \tag{B.16}$$

Filling in the volume accelerations of equations B.14 and B.15 into equation B.16 leads to an expression that relates the volume acceleration of a section of the basilar

membrane to the pressures acting on that section via:

$$\dot{U}_n = \frac{1}{dx\, M_{sn}} \left(p_{n-1} - 2p_n + p_{n+1}\right) \tag{B.17}$$

With expression B.2 for the volume velocity U, equation B.17 can be rewritten into an expression for the basilar membrane acceleration in the form:

$$\ddot{y}_n = \frac{1}{b\, dx^2\, M_{sn}} \left(p_{n-1} - 2p_n + p_{n+1}\right) . \tag{B.18}$$

To complete the pressure balance over a specific section of the basilar membrane and thus the description of the transmission-line model, the series and parallel impedance need to be coupled together. This is done by filling in the basilar-membrane acceleration expression B.18, found from the series impedance Z_s into the pressure equation B.4 derived from the shunt impedance Z_p.

$$p_n - \frac{\omega_{c0}\, M_{p0}}{\omega_{cn}\, M_{sn}\, dx} \left(p_{n-1} - 2p_n + p_{n+1}\right) = \omega_{c0}\, b\, M_{p0} \left[\delta\, \dot{y}_n + \omega_{cn} \left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{cn}}}\right) y_n\right] \tag{B.19}$$

Multiplying both sides of the equation with:

$$\frac{\omega_{cn}\, M_{sn}\, dx^2}{\omega_{c0}\, M_{p0}} , \tag{B.20}$$

leads to:

$$-p_{n-1} + \left[2 + \frac{\omega_{cn}\, M_{sn}\, dx^2}{\omega_{c0}\, M_{p0}}\right] p_n - p_{n+1} = M_{sn}\, dx^2\, b \left[\omega_{cn} \delta\, \dot{y}_n + \omega_{cn}^2 \left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{cn}}}\right) y_n\right] . \tag{B.21}$$

Rewriting this equation into pressure per unit mass of the basilar membrane q instead of into the pressure p, by using expression B.7, leads to the following expression for the transmission line model:

$$-q_{n-1} + \left[2 + \frac{\omega_{cn}\, M_{sn}\, dx^2}{\omega_{c0}\, M_{p0}}\right] q_n - q_{n+1} = \frac{\omega_{cn}\, M_{sn}\, dx^2}{\omega_{c0}\, M_{p0}} \left[\omega_{cn} \delta\, \dot{y}_n + \omega_{cn}^2 \left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{cn}}}\right) y_n\right] . \tag{B.22}$$

The following constants can be defined from expression B.22:

$$A_{sq} = \frac{\omega_{cn}\, M_{sn}\, dx^2}{\omega_{c0}\, M_{p0}}\,, \tag{B.23}$$

and

$$A_{sc} = 2 + A_{sq}\,, \tag{B.24}$$

with

$$[A_{sq}] = [A_{sc}] = \left[\frac{1}{T}\frac{M}{L^5}L^2 T\frac{L^3}{M}\right] = [\,]\,. \tag{B.25}$$

In its simplified form, the describing transmission line equation thus reads:

$$-\,q_{n-1} + A_{sc}\,q_n - q_{n+1} = A_{sq}[D\,\dot{y}_n + S\,y_n]\,. \tag{B.26}$$

B.2 The middle-ear boundary

Applying equation 4.17 and Newton's second law to the middle-ear boundary in Fig.4.1 leads to the following expression at the middle-ear boundary of the transmission line model:

$$p_{stim} - p_0 = -R_{me}\,U_0 - \frac{M_{s0}\,dx}{2}\,\dot{U}_0\,. \tag{B.27}$$

An expression for $\dot{U}_0$ can be found by applying equation 4.3 for the series impedance of the model to the first section in the schematic in Fig.4.1:

$$\dot{U}_0^+ = \frac{2(p_0 - p_1)}{dx\, M_{s0}}\,. \tag{B.28}$$

When applying conservation of mass at node 0, it follows that:

$$U_0 = -U_0^+\,, \tag{B.29}$$

After inserting equation B.29 into equation B.27, and after applying equation B.2 and B.7, the following expression is found for the pressure balance at the middle-ear

boundary:

$$2q_0 - q_1 = \frac{p_{stim}}{M_{p0}\, b} + \frac{R_{me}\, dx}{M_{p0}}\, \dot{y}\,. \tag{B.30}$$

The pressure balance at section 1 of the cochlea will furthermore be described differently than the following sections, and this is due to the acoustical series mass $dxM_{s0}/2$, that was placed between p_0 and p_1 (see Fig.4.1). Applying conservation of mass at section 1 of the schematic leads to the following expression of $\dot{U}_1$:

$$\dot{U}_1 = \frac{1}{M_{s0}\, dx}(2p_0 - 3p_1 + p_2) \tag{B.31}$$

Filling in equation B.31 into equation B.4 leads to the following balance at the first section of the cochlea:

$$-\,2q_0 + A_{sc1}\, q_1 - q_2 = A_{sq1}[D\, \dot{y}_n + S\, y_n] \tag{B.32}$$

where

$$A_{sq1} = \frac{\omega_{c1}\, dx^2\, M_{s0}}{\omega_{c0}\, M_{p0}}\,, \text{ and} \tag{B.33}$$

$$A_{sc1} = A_{sq1} + 3\,. \tag{B.34}$$

B.3 The helicotrema boundary

To describe the last section of the model, an expression for $\ddot{y}$ at the helicotrema boundary was derived. Applying equation 4.3 to the schematic in Fig.4.1 led to the following expression for the series impedance of the basilar membrane:

$$M_{sn} = \frac{-(p_n - p_{n-1})}{dx\, U^+_{n-1}}\,. \tag{B.35}$$

By use of equation B.2, and by applying the conservation of mass principle to node n in the schematic of Fig.4.1:

$$U_n = U^+_{n-1} - U_H\,, \tag{B.36}$$

equation B.35 can be rewritten into:

$$\ddot{y} = \frac{-(p_n - p_{n-1})}{b\, dx^2\, M_{sn}}\,. \tag{B.37}$$

Inserting equation B.37 into equation B.4 yields:

$$p_n + \frac{\omega_{c0}\, M_{p0}}{\omega_{cn}\, dx^2\, M_{sn}}(p_n - p_{n-1}) = b\,\omega_{c0}\, M_{p0}\left[\delta\dot{y}_n + \omega_{cn}\left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{cn}}}\right) y_n\right]\,. \tag{B.38}$$

Multiplying both sides of the equation with:

$$\frac{\omega_{cn}\, dx^2\, M_{sn}}{\omega_{c0}\, M_{p0}} \tag{B.39}$$

leads to the following pressure balance at the helicotrema boundary of the transmission line model:

$$\left[1 + \frac{\omega_{cn}\, dx^2 M_{sn}}{\omega_{c0} M_{p0}}\right]\, p_n - p_{n-1} = b\,\omega_{c0}\, M_{p0}\left[\delta\dot{y}_n + \omega_{cn}\left(1 + \rho \mathrm{e}^{\frac{-j\omega\psi}{\omega_{cn}}}\right) y_n\right]\,. \tag{B.40}$$

Rewriting equation B.40 to the format of the general transmission line equation B.26 gives the following expression for the transmission line model equation at the helicotrema boundary:

$$A_{hc}\, q_n - q_{n-1} = A_{sq}\left[D\, y_n + S\, y_n\right]\,, \tag{B.41}$$

where:

$$A_{hc} = 1 + A_{sq} \tag{B.42}$$

and where A_{sq}, D and S are respectively defined in equation B.23, B.9 and B.10.

Bibliography

Békésy, G. von (1960). Experiments in hearing. McGraw-Hill Book Company, New York.

Boer, E. de (1975). Synthetic whole-nerve action potentials for the cat, J Acoust Soc Am, 58, 1030–1045.

Boer, E. de (1980). Auditory physics. Physical principles in hearing theory. I. Phys Rpt, 62, 88–174.

Brownell, W. E. and Bader, C. E. and Bertrand, D. and deRibaupierre, Y. (1985). Evoked mechanical responses of isolated outer hair cells. Science, 227, 194–196.

Carney, L. H. and McDuffy, M. J. and Shekhter, I. (1999). Frequency glides in the impulse responses of auditory-nerve fibers. J Acoust Soc Am, 105, 2384–2391.

Carvalho, S. and Büki, B. and Bonfils, P. and Avan, P. (2003). Effect of click intensity on click-evoked otoacoustic emission waveforms: implications for the origin of emissions. Hear Res, 175, 215–225.

Davis H. (1983). An active process in cochlear mechanics. Hear Res, 9, 79–90.

Diependaal, R. J. and Duifhuis, H. and Hoogstraten, H. W. and Viergever, M. A. (1987). Numerical methods for solving 1-dimensional cochlear models in the time domain. J Acoust Soc Am, 82 (5), 1655-1666.

Duifhuis, H. and Hoogstraten, H .W. and Netten, S. M. van and Diependaal, R. J. and Bialek, W. (1986). Modelling the cochlear partition with coupled Van der Pol oscillators. In: Allen, J. B. and Hall, J. L. and Hubbard, A. E. and Neely S. T. and Tubis, A. Peripheral auditory mechanims. Springer-Verlag, New York, 290–297.

Eguiluz, V. M. and Ospeck, M. and Choe, Y. and Hudspeth, A. J. and Magnasco, M. O. (2000). Essential nonlinearities in hearing. Phys Rev Lett, 84 (22), 5232–5235.

Elliott, S. J. and Ku, E. M. and Lineton, B. (2007). A state space model for cochlear mechanics. J Acoust Soc Am, 87, 2592–2605.

Geisler C. D. (1991). A cochlear model using feedback from motile outer hair cells. Hear Res, 54, 105–117.

Gentle, J. E. (1998). Gaussian Elimination 3.1. In: Numerical Linear Algebra for Applications in Statistics, Springer-Verlag, Berlin, 87–91.

Glasberg, B. R. and Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. J Acoust Soc Am, 79, 1020-1033.

Glasberg, B. R. and Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. Hear Res, 47 (1–2), 103–138.

Gold, T. (1948). Hearing. II. The physical basis of the action of the cochlea. Proceedings of the Royal Society of London, 135 (881), 492–498.

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

Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. J Acoust Soc Am, 33, 1344-1356.

Harte, J. M. and Elliot, S. J. and Kapadia, S. and Lutman, M. E. (2005). Dynamic nonlinear cochlear model predictions of click-evoked otoacoustic emission suppression. Hear Res, 207, 99–109.

Hengel, P.W.J. van (1996). Emissions from cochlear modelling. PhD thesis, Rijksuniversiteit Groningen.

Hine, J. E. and Thornton, A. R. D. (2002). Temporal nonlinearity revealed by transient evoked otoacoustic emissions recorded to trains of multiple clicks. Hear Res, 165, 128–141.

Hudspeth, A. J. and Gillespie, P. G (1994). Pulling springs to tune transduction: adaptation by hair cells. Neuron, 12, 1–9.

Kalluri, R. and Shera, C. A. (2007). Near equivalence of human click-evoked and stimulus-frequency otoacoustic emissions. J Acoust Soc Am, 121 (4), 2097–2110.

Kapadia, S. and Lutman, M. E. (2000a). Nonlinear temporal interactions in click-evoked otoacoustic emissions. I. Assumed model and polarity-symmetry. Hear Res, 146 (1-2), 89–100.

Kapadia, S. and Lutman, M. E. (2000b). Nonlinear temporal interactions in click-evoked otoacoustic emissions. II. Experimental data, Hear Res, 146 (1-2), 101–120.

Kapadia, S. and Lutman, M. E. and Palmer, A. R. (2005). Transducer hysteresis contributes to "'stimulus artifact"' in the measurement of click-evoked otoacoustic emissions (L) J. Acoust. Soc. Am., 118(2), 620–622.

Keefe, D. H. and Ellison, J. C. and Fitzpatrick, D. F. and Gorga, M. P. (2007). Two-tone suppression of stimulus frequency otoacoustic emissions. J Acoust Soc Am, 123 (3), 1479–1494.

Kemp, D. T. and Martin J. A. (1976). Active resonant systems in audition. Abstracts of XIII international congress of audiology Firenze 1976. International society of audiology Geneva, 64–65.

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

Kemp, D. T. and Chum, R. A. (1980). Properties of the generator of stimulated otoacoustic emissions. Hear Res, 2, 213–232.

Kemp, D. T. and Bray, P. and Alexander, L. and Brown, M. A. (1986). Acoustic emission cochleography–Practical aspects. Scand. Audiol. Suppl., 25, 71–82.

Kemp, D. T. (2007). Otoacoustic Emissions: Concepts and Origins. In Manley, A. and Fay, R. R. and Popper, A. N. (Ed.) Active Processes and Otoacoustic Emissions in Hearing (pp. 1–38). New York: Springer.

Z. Kevanishvili and G. Tietze and H. Gobsch (1996). Effects of the conditioning click on click-evoked otoacoustic emissions. Scand. Audiol., 25(3), 161–166.

Kim, D. O. and Molnar, C. E. and Matthews, J. W. (1980b). An active cochlear model with negative damping in the partition: comparisons with Rhode's ante- and post-mortem observations. In: van den Brink, G. and Bilsen, F. A. Psychophysical, physiological and behavioral studies in hearing. Delft university press, Delft, 7–14.

Kirk, D. L. and Yates, G. K. (1996). Frequency tuning and acoustic enhancement of electrically evoked otoacoustic emissions in the guinea pig cochlea. J Acoust Soc Am, 100 (6), 3714–3725.

Kirk, D. L. and Yates, G. K. (1998). Enhancement of electrically evoked oto-acoustic emissions associated with low-frequency stimulus bias of the basilar membrane towards scala vestibuli. J Acoust Soc Am, 104 (3), 1544–1554.

Lina-Granade, G. and Collet, L. (1995). Effect of interstimulus interval on evoked otoacoustic emissions. Hear Res, 87, 55–61.

Lineton, B. and Kuponiyi, O. and Thornton, A. R. D. (2008). The effect of stimulus rate and time-separation on Voltera slices of otoacoustic emissions. Hear Res, 239 (1–2), 34–53.

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. J Acoust Soc Am, 63, 442–255.

Lineton, B. and Wildgoose, C. M. B. (2009). Comparing two proposed measures of cochlear mechanical filter bandwidth based on stimulus frequency otoacoustic emissions. J Acoust Soc Am, 125 (3), 1558–1566.

Liu, Y. and Neely, S. T. (2009). Outer hair cell electromechanical properties in a nonlinear piezoelectric model. J Acoust Soc Am, 126 (2), 751–761.

Lopez-Poveda, E. A. and Meddis, R. (2001). A human nonlinear cochlear filterbank. J Acoust Soc Am, 110 (6), 3107–3118.

Lopez-Poveda, E. A. and Plack, C. J. and Meddis, R. (2003). Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing. J Acoust Soc Am, 113 (2), 951–960.

Lynch, T. J. and Nedzelnitsky, V. and Peak, W. T. (1982). Input impedance of the cochlea in cat. J Acoust Soc Am, 72 (1), 108–130.

Martin, P. (2007). Active Hair-Bundle Motility of the Hair Cells of Vestibular and Auditory Organs. In Manley, A. and Fay, R. R. and Popper, A. N. (Ed.) Active Processes and Otoacoustic Emissions in Hearing (pp. 102–112). New York: Springer.

Moleti, A. and Paternoster, N. and Bertaccini, D. and Sisto, R. and Sanjust, F. (2009). Otoacoustic emissions in time-domain solutions of nonlinear non-local cochlear models. J Acoust Soc Am, 126 (5), 2425–2436.

Moore, B. C. J. and Glasberg, B. R. and Stone, M. A. (1984). Refining the measurement of psychophysical tuning curves. J Acoust Soc Am, 76, 1057–1066.

Moore, B. C. J. and Stainsby, T. H. and Alcantara, J. I. and Kuhnel, V. (2004). The effect on speech intelligibility of varying compression time constants in a digital hearing aid. Int J Aud, 43 (7), 399–409.

Mountain, D. C. and Hubbard, A. E. (1989). Rapid force production in the cochlea. Hear Res, 42, 195–202.

Nedzelnitsky, V. (1980). Sound pressures in the basal turn of the cat cochlea. J Acoust Soc Am, 69, 1676–1689.

Neely S. T. (1993). A model of cochlear mechanics with outer hair cell motility. J Acoust Soc Am, 94 (1), 137–146.

Netten, S. M. van and Duifhuis, H. (1983). Modelling an active, nonlinear cochlea. In: Boer, E. de and Viergever, M. A.. Mechanics of Hearing, Martinus Nijhoff Publishers, Den Haag and Delft University Press, Delft, 143–151.

Neumann, J. and Uppenkamp, S. and Kollmeier, B. (1997). Interaction of otoacoustic emissions with additional tones: suppression or synchronisation? Hear Res, 103, 19–27.

Pigasse, G. (2008). Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials. PhD thesis, Technical University of Denmark.

Prieve, B. A. and Falter, S. R. (1995). COAEs and SSOAEs in Adults with Increased Age. Ear Hear, 16 (5), 521–528.

Prieve, B. A. and Fitzgerald, T. S. and Schulte, L. E. (1997). Basic characteristics of click-evoked otoacoustic emissions in infants and children. J Acoust Soc Am, 102 (5), 2860–2870.

Probst, R. and Coats, A. C. and Martin, G. K. and Lonsbury-Martin, B. L. (1986). Spontaneous, click-, and toneburst-evoked otoacoustic emissions from normal ears. Hear Res, 21, 261–275.

Probst, R. and Lonsbury-Martin, B. L. and Martin, G. K. (1991). A review of otoacoustic emissions. J Acoust Soc Am, 89 (5), 2027–2067.

Puria, S. (2003). Measurements of human middle ear forward and reverse acoustics: Implications for otoacoustic emissions. J Acoust Soc Am, 113 (5), 2773–2789.

Puria, S. and Allen, J. B. (1991). A parametric study of cochlear input impedance. J Acoust Soc Am, 89 (1), 287–309.

Recio, A. and Rhode, W. S. (2000). Basilar membrane responses to broadband stimuli. J Acoust Soc Am, 108 (5), 2281–2298.

Ren, T. (2002). Longitudinal pattern of basilar membrane vibration in the sensitive cochlea. PNAS, 99 (26), 17101-17106.

Ren, T. and Nuttall, A. L. (1999). Comment on "Enhancement of the transient-evoked otoacoustic emission produced by the addition of a pure tone in the guinea pig" [J.Acoust.Soc.Am.104,344-349(1998)]. J Acoust Soc Am, 105(2), 919–921.

Ren, T. and Nuttall, A. L. (2001). Basilar membrane vibration in the basal turn of the sensitive gerbil cochlea. Hear Res, 151, 48–60.

Rhode, W. S. (1971). Observations of vibration of basilar membrane in squirrel monkeys using mossbauer technique. J Acoust Soc Am, 49 (4), 1218–1231.

Rhode, W. S. and Cooper, N. P. (1996). Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo, Aud Neurosci, 3, 101Ű121.

Rhode, W. S. and Recio, A. (2000). Study of mechanical motions in the basal region of the chinchilla cochlea. J Acoust Soc Am, 107 (6), 3317–3331.

Robles, L. and Ruggero, M. A. and Rich, N. C. (1986). Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases. J Acoust Soc Am, 80, 1364–1374.

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

Schairer, K. S. and Ellison, J. C. and Fitzpatrick, D. and 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 (2), 901–914.

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

Shera, C. A. (1992). Listening to the ear. Ph.D. thesis. California Institute of Technology.

Shera, C. A. (2001). Frequency glides in click responses of the basilar membrane and auditory nerve: Their scaling behavior and origin in traveling-wave dispersion. J Acoust Soc Am, 109 (5), 2023–2034.

Shera, C. A. (2001). Intensity-invariance of fine time structure in basilar-membrane click responses: Implications for cochlear mechanics. J Acoust Soc Am, 110 (1), 332–348.

Shera, C. A. (2003). Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves. J Acoust Soc Am, 114 (1), 244–262.

Shera, C. A. (2007). Laser amplification with a twist: Traveling-wave propagation and gain functions from throughout the cochlea. J Acoust Soc Am, 122 (2), 2738–2758.

Shera, C. A. and Zweig, G. (1991). A symmetry suppresses the cochlear catastrophe. J Acoust Soc Am, 89 (3), 1276 -1289.

Shera, C. A. and Guinan, J. J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs. J Acoust Soc Am, 105, 782–798.

Shera, C. A. and Guinan, J. J. Jr and Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioural measurements. Proc Natl Acad Sci USA, 99, 3318–3323.

Shera, C. A. and Guinan, J. J. (2007), Mechanisms of Mammalian Otoacoustic Emission. In Manley, A. and Fay, R. R. and Popper, A. N. (Ed.) Active Processes and Otoacoustic Emissions in Hearing (pp. 306–342). New York: Springer.

Shera, C. A. and Cooper, N. P. (2009), Wave interference patterns on the basilar membrane: testing models of OAE propagation. In proceedings of the 32rd annual midwinter research meeting of the association for research in otolaryngology. Feb 14–19 2009, Baltimore, Maryland.

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

Tavartkiladze, G. A. and Frolenkov, G. I. and Kruglov, A. V. and Artamasov, S. V. (1994). Ipsilateral suppression effects on transient evoked otoacoustic emission. Br J Audiol, 28 (4-5), 193–204.

Thornton, A. R. D and Lineton, B. and Baker, V. J. and Slaven, A. (2006). Nonlinear properties of otoacoustic emissions in normal and impaired hearing. Hear Res, 219 (1–2), 56–65.

Viergever, M. A. (1980). Mechanics of the inner ear: A mathematical approach. Delft U.P., Delft

Verhulst, S. and Harte, J. M. and Dau, T. (2008). Temporal suppression and augmentation of click-evoked otoacoustic emissions. Hear. Res., 246, 23–35.

Verhulst, S. and Harte, J. M. and Dau, T. (2011). Temporal suppression of the click-evoked otoacoustic emissions level-curve. J Acoust Soc Am, 129 (2) in press.

Wilson, J. P. (1980). Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus. Hear Res, 2, 233-252.

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

Withnell, R. H. and Hazlewood, C. and Knowlton A. (2008). Reconciling the origin of the transient evoked otoacoustic emission in humans. J Acoust Soc Am, 123 (1), 212–221.

Xue, S. and Mountain, D. C. and Hubbard, A. E. (1993). Acoustic enhancement of electrically-evoked otoacoustic emissions reflects basilar membrane tuning: Experimental results. Hear res, 70, 121–126.

Yates, G. K. and Withnell, R. H. (1999). Reply to "Comment on 'Enhancement of the transient-evoked otoacoustic emission produced by the addition of a pure tone in the guinea pig'" [J.Acoust.Soc.Am.105, 919-921(1999)]. J Acoust Soc Am, 105(2), 922–924.

Zweig, G. (1976). Basilar membrane motion. In: Cold spring harbor symposia on Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, Vol. XL, 619-Ű633.

Zweig, G. (1991). Finding the impedance of the organ of Corti. J Acoust Soc Am, 89 (3), 1229-1254.

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

Contributions to Hearing Research

Vol. 1: *Gilles Pigasse*, Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials, Dec. 2008.

Vol. 2: *Olaf Strelcyk*, Peripheral auditory processing and speech reception in impaired hearing, Jun. 2009.

Vol. 3: *Eric R. Thompson*, Characterizing binaural processing of amplitude-modulated sounds, Aug. 2009.

Vol. 4: *Tobias Piechowiak*, Spectro-temporal analysis of complex sounds in the human auditory system, Sept. 2009.

Vol. 5: *Jens Bo Nielsen*, Assessment of speech intelligibility in background noise and reverberation, Dec. 2009.

Vol. 6: *Helen Connor Sørensen*, Hearing aid amplification at soft input levels, Jan. 2010.

Vol. 7: *Morten Løve Jepsen*, Modeling auditory processing and speech perception in hearing-impaired listeners, May 2010

Vol. 8: *Sarah Verhulst*, Characterizing and modeling dynamic processes in the cochlea using otoacoustic emissions, Jun. 2010

Vol. 9: *Sylvain Favrot*, A loudspeaker-based room auralization system for auditory research, Jun. 2010

An important characteristic of human hearing is that it amplifies weak sounds while attenuating louder ones. This gain transformation takes place in the inner ear (i.e., cochlea), and is responsible for a compressive relation between the level of the presented and perceived sound. The cochlear gain mechanism is essential for our hearing and degrades when hearing impairment develops. A comprehensive understanding of the gain involved in the intact human cochlea is crucial, as hearing instruments try to compensate for the loss in cochlear gain caused by hearing damage. This thesis investigates dynamic, or time-dependent, properties of cochlear gain. A time scale from 0 to 10 ms is considered to ensure that cochlear processing is investigated without including influences from higher stages in the brain. The results provide insight into how e.g. onsets of sounds are processed by the intact human system.

Centre for Applied Hearing Research
Department of Electrical Engineering

Ørsteds Plads
Building 352
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 38 00
Fax: (+45) 45 93 16 34
www.elektro.dtu.dk

ISBN 978-87-92465-24-5