

# Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials

Ph.D. thesis by  
Gilles Pigasse



Technical University of Denmark  
2008

Copyright © Gilles Pigasse, 2008  
ISBN 978-87-911-8489-5  
Printed in Denmark by Hillerød Grafisk.

# Contents

<b>Preface</b>	<b>vii</b>
<b>Abstract</b>	<b>ix</b>
<b>List of abbreviations and symbols</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Project motivation and objectives . . . . .	2
1.2 Contributions of the thesis . . . . .	4
<b>2 Background and theory</b>	<b>5</b>
2.1 Otoacoustic emissions . . . . .	5
2.1.1 Generation mechanisms . . . . .	8
2.1.2 Effect of external factors on OAEs . . . . .	9
2.1.3 Clinical usage of OAEs . . . . .	10
2.2 Auditory evoked potentials . . . . .	10
2.2.1 Auditory brainstem responses . . . . .	14
2.2.2 Auditory steady-state responses . . . . .	17
2.2.3 Comparison between ASSRs and ABRs . . . . .	20
2.3 Summary . . . . .	21
<b>3 Individual cochlear delay estimates using otoacoustic emissions</b>	<b>23</b>
3.1 Introduction . . . . .	24
3.2 Methods . . . . .	25

---

3.2.1	Subjects . . . . .	25
3.2.2	Stimulus generation and response measurement . . . . .	25
3.2.3	Off-line data analysis . . . . .	28
3.3	Results . . . . .	35
3.3.1	Effect of frequency on TBOAE latency . . . . .	35
3.3.2	Intra-subject variability . . . . .	38
3.3.3	Inter-subject variability . . . . .	40
3.3.4	Effect of the noise floor on OAE latency . . . . .	41
3.3.5	Comparison of TBOAE latency with previous studies . . . . .	43
3.4	Discussion . . . . .	44
3.4.1	Conclusion . . . . .	47
<b>4</b>	<b>Individual estimates of brainstem response delays</b>	<b>49</b>
4.1	Introduction . . . . .	49
4.2	Methods . . . . .	51
4.2.1	Subjects . . . . .	51
4.2.2	Stimuli . . . . .	51
4.2.3	Procedure . . . . .	52
4.2.4	TBABR latency and interpeak delays . . . . .	53
4.3	Results . . . . .	55
4.3.1	wave-V latency . . . . .	55
4.3.2	Intra- and inter-subject variability . . . . .	57
4.3.3	Interpeak intervals . . . . .	59
4.4	Discussion . . . . .	60
4.4.1	Rise time difference . . . . .	62
4.4.2	Level difference . . . . .	62
4.4.3	Subject gender difference . . . . .	63
4.4.4	Summary . . . . .	63
<b>5</b>	<b>Comparison of OAE and ABR estimates of cochlear delay</b>	<b>65</b>
5.1	Introduction . . . . .	66
5.1.1	Acoustic wave hypothesis . . . . .	66
5.1.2	Coherent Reflection Filtering theory . . . . .	66



---

5.1.3	Signal-front hypothesis . . . . .	67
5.2	Results . . . . .	68
5.2.1	Intra-subject variability . . . . .	70
5.2.2	Inter-subject variability . . . . .	70
5.2.3	Acoustic wave hypothesis . . . . .	71
5.2.4	CRF theory . . . . .	72
5.2.5	Signal-front hypothesis . . . . .	75
5.3	Discussion . . . . .	77
5.3.1	Separating the analysis below and above 2 kHz . . . . .	78
5.3.2	Scaling symmetry . . . . .	80
5.4	Conclusion . . . . .	82
<b>6</b>	<b>Individual steady-state response delays</b>	<b>85</b>
6.1	Introduction . . . . .	85
6.2	Methods . . . . .	87
6.2.1	Subjects . . . . .	87
6.2.2	Stimuli and recording procedure . . . . .	88
6.2.3	Response detection . . . . .	89
6.2.4	ASSR latency . . . . .	90
6.3	Results . . . . .	91
6.3.1	Latency-frequency functions . . . . .	91
6.3.2	Intra-subject variability . . . . .	93
6.3.3	Comparison with previous studies . . . . .	93
6.4	Discussion . . . . .	95
6.4.1	Comparison with ABR . . . . .	95
6.4.2	Gender difference . . . . .	98
6.5	Conclusion . . . . .	99
<b>7</b>	<b>Chirp-evoked otoacoustic emissions</b>	<b>101</b>
7.1	Introduction . . . . .	101
7.2	Methods . . . . .	104
7.2.1	Subjects and procedure . . . . .	104
7.2.2	Stimuli . . . . .	105

7.2.3	Time-frequency distribution . . . . .	106
7.3	Results . . . . .	108
7.3.1	OAE time series . . . . .	108
7.3.2	Time-frequency analysis . . . . .	111
7.4	Discussion and conclusion . . . . .	114
<b>8</b>	<b>Overall discussion</b>	<b>115</b>
8.1	Summary of the main results . . . . .	115
8.2	Discussion . . . . .	117
8.3	Implications and future work . . . . .	119
	<b>Bibliography</b>	<b>123</b>
	<b>Appendix</b>	<b>139</b>
<b>A</b>	<b>Tone bursts: time and frequency plots</b>	<b>141</b>
<b>B</b>	<b>Stimuli calibration</b>	<b>145</b>
<b>C</b>	<b>Mathematical tools</b>	<b>147</b>
C.1	Hilbert transform . . . . .	147
C.2	The least-squares fit time-domain filtering . . . . .	148
C.3	Short-time correlation coefficient . . . . .	148
C.4	Pearson product moment correlation coefficient . . . . .	149
<b>D</b>	<b>TBOAE and TBABR latencies</b>	<b>151</b>
<b>E</b>	<b>Effect of tone burst rise time on wave-V latency</b>	<b>153</b>
<b>F</b>	<b>Acoustic wave theory results</b>	<b>157</b>
<b>G</b>	<b>CRF results</b>	<b>159</b>
<b>H</b>	<b>Front delay results</b>	<b>161</b>
<b>I</b>	<b>Individual ASSR latencies</b>	<b>163</b>

# Preface

This thesis was submitted to the Technical University of Denmark (DTU) as partial and final fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.) in Electronics and Communication. The work presented in this thesis was completed between October 1, 2004 and July 31, 2008 at the Centre for Applied Hearing Research (CAHR), Department of Electrical Engineering, DTU and at the Institute of Sound and Vibration Research (ISVR), Southampton University, England. The work was done under the supervision of Prof. Torsten Dau and Dr. James Harte. A public defense was held on December 3, 2008. The project was jointly funded by GN ReSound A/S, Oticon A/S and Widex A/S and the external stay at ISVR was supported by a Marie Curie Research Fellowship via the European Doctorate in Sound and Vibration Studies. Ethical considerations for conducting the experiments at DTU were made with respect to the Scientific Ethics Committee of Copenhagen County, file number KA 04159g, experiments made at ISVR were approved by the departmental Safety and Ethics Committee of this institute. Every chapter presenting experimental results contains, prior to its introduction, an abstract, which is intended for the reader to have a rapid overview of the work presented herein. A list of the main abbreviations and symbols used in this thesis is available before the table of content. The appendices contain plots, tables and explanations that are not crucial to understand the main part, but can be referred to for more details about the experimental work.

The work described in this thesis could not have been completed without the help and support of many people, I would like to thank my two supervisors Torsten and James for guiding me through this project, especially during times of immense doubt.

Thanks also to Manfred Mauermann and his colleagues at the Haus Des Hörens and the University of Oldenburg, for welcoming me in their group during three months, making interesting suggestions and for a fruitful collaboration when writing the OAE software. Thanks to Steve Bell and his colleagues from the Hearing and Balance Centre at Southampton University for a pleasant and useful stay. Thanks to my colleagues in CAHR, especially Tobias Piechowiak and Jens Bo Nielsen for keeping a nice atmosphere in the office, Iris Arweiler and Brent Kirkwood for proofreading part of this work and providing useful advice. Thanks also to the 55 (or so) test persons for their patience and their will to sleep for hours during the experiments, even if not all data could be used at the end.

---

Gilles Pigasse

January 7, 2009

# Abstract

A great deal of the processing of incoming sounds to the auditory system occurs within the cochlear. The organ of Corti within the cochlea has differing mechanical properties along its length that broadly gives rise to frequency selectivity. Its stiffness is at maximum at the base and decreases towards the apex, resulting in locally resonant behaviour. This means high frequencies have maximal response at the base and low frequencies at the apex. The wave travelling along the basilar membrane has a longer travel time for low-frequency stimulus than for high-frequency stimulus. The intrinsic relation between frequency and travel time in the cochlea defines the cochlear delay. This delay is directly associated with the signal analysis occurring in the inner ear and is therefore of primary interest to get a better knowledge of this organ. It is possible to estimate the cochlear delay by direct and invasive techniques, but these disrupt the normal functioning of the cochlea and are usually conducted in animals. In order to obtain an estimate of the cochlear delay that is closer to the normally functioning human cochlea, the present project investigates non-invasive methods in normal hearing adults. These methods include: otoacoustic emissions (OAEs), auditory brainstem responses (ABRs) and auditory steady-state responses (ASSRs).

A comparison between the three methods was made across and within subjects, in order to highlight the impact of inter-subject variability on the cochlear delay estimates. The estimates of the cochlear delay obtained with OAEs, ABRs and ASSRs were in good agreement with previously reported studies.

The comparison between OAE and ABR latency estimates was made over a broader range of frequencies (0.5-8 kHz), compared to previous studies. Below about 2 kHz the OAE delay is twice the cochlear delay, as if the travelling wave went

back and forth in the cochlea, as predicted in current theories of OAE generation. This relation, however, does not hold for higher frequencies, calling into question the physical relation between OAE and ABR delay estimates. The comparison between ABR and ASSR latency estimates demonstrated similar rates of latency decrease as a function of frequency. It was further concluded, in this thesis, that OAE measurements are the most appropriate to estimate cochlear delays, since they had the best repeatability and the shortest recording time. Preliminary results are also given for an experiment using stimuli designed to compensate for OAE delays. These were designed to try and reproduce the success of similar stimuli now used routinely to improve ABR signal-to-noise ratio.

# Resumé

En stor del af lydbehandlingen i det auditive system sker i sneglen. Sneglens cortiske organ ændrer mekaniske egenskaber i længderetningen, og denne ændring forklarer hovedparten af den frekvensadskillelse, som finder sted. Basilarmembranens stivhed er maximal ved basen og falder mod apex, hvilket resulterer i lokale resonanser for forskellige frekvenser langs hele membranen. Høje frekvenser giver anledning til maximalt udslag ved basen; for lave frekvenser sker det ved apex. Bølgen, som forplanter sig langs basilarmembranen, bevæger sig langsommere for lav-frekvent stimulus end for høj-frekvent stimulus. Udberedelsestiden fra basen til resonansstedet for en bestemt frekvens er det såkaldte *cochlear delay* for denne frekvens. Denne tid er tæt forbundet med signalanalysen i det indre øre og er derfor af stor interesse i udforskningen af dette organ. Det er muligt at estimere *cochlear delay* med direkte og invasive teknikker, men disse forstyrrer ørets normale funktion og udføres normalt kun på dyr. For at estimere *cochlear delay* i et normalt fungerende menneskeligt øre udforsker dette projekt ikke-invasive metoder med normalthørende voksne. Disse metoder inkluderer otoacoustic emissions (OAEs), auditory brainstem responses (ABRs) og auditory steady-state responses (ASSRs).

En sammenligning mellem de tre metoder blev gennemført specielt med henblik på at undersøge variationen af forsinkelsesestimaterne ved gentagne målinger med samme forsøgsperson. Estimaterne, som blev målt med OAEs, ABRs og ASSRs var i god overensstemmelse med tidligere forskningsresultater.

Sammenligningen mellem OAE- og ABR-estimaterne blev gennemført over et bredere frekvensområde (0.5-8 kHz) end i tidligere studier. Under ca. 2 kHz er OAE forsinkelsen to gange *cochlear delay*'et, hvilket antyder, at bølgen har forplantet

sig både frem og tilbage langs basilarmembranen. Dette er i overensstemmelse med de aktuelle OAE-teorier. Sammenhængen eksisterer dog ikke for højere frekvenser, og det rejser spørgsmålet om den fysiske sammenhæng mellem OAE- og ABR-forsinkelser. En sammenligning mellem ABR og ASSR forsinkelsesestimer udviser det samme fald, når frekvensen stiger. I denne afhandling bliver det også konkluderet, at OAE-målinger er de bedste til at estimere *cochlear delay*'er, fordi de har den bedste reproducerbarhed og den korteste måletid. Præliminære resultater fra et eksperiment med stimuli, der kompenserer for OAE-udbredelsestiden, præsenteres også. Disse er designet på baggrund af lignende stimuli, som nu rutinemæssigt bruges til at forbedre ABR signal/støj-forholdet.



# List of abbreviations and symbols

ABR	Auditory brainstem response
AEP	Auditory evoked potential
ANOVA	Analysis of variance
ASSR	Auditory steady-state response
BM	Basilar membrane
CF	Characteristic frequency
CI	Cochlear implants
CRF	Coherent reflection filtering
DPOAE	Distortion product OAE
dB HL	dB hearing level
dB nHL	dB normal hearing level
dB peSPL	dB peak-equivalent sound pressure level
dB SL	dB sensation level
EC	Ear canal
ECochG	Electrocochleography
IHC	Inner hair cell
LAEP	Late AEP
MLR	Middle latency response
OAE	Otoacoustic emission
OHC	Outer hair cell

---

PAMR	Post-auricular-muscle response
SFOAE	Stimulus frequency OAE
SNR	Signal-to-noise ratio
SOAE	Spontaneous OAE
Std	Standard deviation
STCC	Short-time correlation coefficient
TB	Tone burst
TBABR	Tone-burst ABR
$\hat{T}_{BM}^{(ABR)}$	BM latency estimated from ABR
TBOAE	Tone-burst OAE
TEOAE	Transient evoked OAE
$\tau_{filter}$	Cochlear filter build-up time
$\tau_{front}$	BM front delay
$\tau_{neural}$	Neural conduction time
$\tau_{OAE}$	OAE latency estimate
$\tau_{synaptic}$	Synaptic delay
$\tau_{transport}$	Cochlear transport time
TW	Travelling wave
$\tau_{wave\ v}$	Wave-V latency

# Introduction

---

Every sound from our environment invariably travels through the same path before it is perceived. This auditory pathway is composed of different stages with specific functions. The outer ear, made of the pinna and ear canal (EC), collects airborne sounds, the middle ear (auditory ossicles) transmits the sound to the inner ear (cochlea) which converts it into nerve impulses. The auditory cortex then analyses the neuronal activity as being a specific sound. Many aspects of hearing can be understood by looking closely at the different properties of the cochlea, such as amplification, frequency analysis or masking. The tonotopic organization of the cochlea is such that highest frequencies are processed at the base of the basilar membrane (BM) and the lowest frequencies are processed at the apex due to a stiffness gradient. This difference in location induces a time difference between the processing of high and low frequencies. The intrinsic relation between frequency and travel time in the cochlea defines the cochlear delay that can provide valuable knowledge about the functioning of the inner ear. It is, in principle, possible to measure directly the BM vibration by laser Doppler velocimetry but every attempt to open the cochlea affects its normal functioning (Dong and Cooper, 2006). Besides, invasive approaches are only used in animals or human cadavers, which only provides an approximation of the normally functioning human cochlea. Despite the difficulty of accessing the cochlea directly and non-invasively, various measurement techniques have been developed to record its activity; for instance electrocochleography (ECochG), where an electrode can be placed in the ear canal to measure the electric potentials generated in the inner ear or recording otoacoustic emissions (OAEs). OAEs are a side effect of the amplification provided by an active process in the cochlea and can be a good indication

of the mechanisms of the inner ear. The cochlear activity is also reflected at higher stages of the human auditory pathway and it is possible to measure it indirectly by placing electrodes on the scalp of a subject. The latency of these auditory evoked potentials (AEP) can be used to estimate the cochlear delay. When measuring cochlear delays, the propagation time in the outer and middle ear is often neglected because it is very short and not frequency dependent, unlike the cochlear delay.

## 1.1 Project motivation and objectives

The aforementioned measurement techniques are already used clinically as diagnostic tools to assess hearing impairment. The diagnosis of cochlear dysfunction is usually based on the detection of the signal (ECochG, OAE), because these signals are an epiphenomenon of the normally functioning cochlea. And, for example, auditory nerve tumors can be detected in case of abnormal or missing AEPs. The cochlear delay derived from OAE or AEP measures is actually not used as a diagnostic tools, maybe because of the time needed to obtain it. The work presented in this thesis investigates different non-invasive methods to estimate the cochlear delay in normal-hearing adults. One of the objectives is to compare three methods and conclude which performs best under given conditions. The three methods are one OAE measurement and two AEP measurements, namely auditory brainstem responses (ABRs) and auditory steady-state responses (ASSRs). An important aim is to obtain information about human cochlear processing and especially about the place on the BM where individual frequencies are filtered. The generation mechanisms of recorded signals can also be compared on the basis of these techniques. It is, to date, not clear if OAEs are generated at the same place as the nerve impulses responsible for the AEP. The focus of this thesis is also on individual measurements, it is important to take into account the inherent differences between subjects since their different cochlear length and state have an impact on the cochlear delay. Comparing individual measures avoids the rough estimation provided by mean data, as it has been the case in previous research. Such an approach is also directly transposable to hearing impaired listeners, for whom an individual delay estimate might be beneficial.

**Chapter 2** gives the knowledge necessary to understand the content of this thesis. This chapter focuses on OAE, ABR and ASSR, their generation mechanism(s), their measurement methods, and their expected latencies.

**Chapter 3** describes the experiment carried out to measure OAEs when tone-bursts are used as stimuli. This chapter also investigates the effect of noise on OAE latency as well as the inter- and intra-subject variabilities of OAE delay. A comparison with previously published data shows how reliable the present results are.

**Chapter 4** presents the experimental work done to measure individual estimates of ABR latencies. Similarly to the OAE experiment, the inter- and intra-subject variabilities are investigated. It is also shown how individual cochlear delays can be derived from ABR delays. In this chapter, the comparison with previously published research is extended and the reason for the differences between studies discussed.

**Chapter 5** is the key chapter of this thesis. It combines the results from the preceding two chapters and puts in parallel, for each subject, the cochlear delay estimated from ABR and OAE. Different theories about OAE generation are presented and confronted with the present results. Conclusions are drawn about the cochlear mechanisms involved in OAE and ABR generation.

**Chapter 6** describes the experiment with which ASSR were recorded. Similarly to the two previous measurement methods, inter- and intra-subject variabilities are investigated. Unfortunately, this experiment could not involve the same subjects as the previous experiments. Hence, there is no direct comparison, but the ASSR latency estimates are averaged and compared with the mean ABR latency estimates. From this comparison, new evidence was found about the generation place of the ASSR.

**Chapter 7** presents a direct application of the data collected in the previous chapters. A short experiment is described in which OAEs are evoked by clicks and chirps. The chirps are generated based on the frequency-latency function obtained in Chap. 3 in order to compensate for the OAE travelling time. The goal is to synchronise the cochlear responses when they reach the ear canal.

**Chapter 8** summarizes the main findings of this thesis and evaluates the comparison between the three methods. This chapter also highlights the implications of the present findings for cochlear modelling and suggests work that could be carried out in the future.

## 1.2 Contributions of the thesis

The main contribution of this work is the comparison of three techniques to measure cochlear delays in humans. The comparison is done for the same subjects between OAE and ABR and is based on average data between ABR and ASSR. Besides comparing results obtained in the present study with results previously found, each experimental study investigates new questions and makes clear contributions to the field, such as:

- a new OAE detection method,
- the study of the effect of noise on OAE latency,
- the comparison, on an individual basis, between OAE and ABR latency estimates,
- a broader frequency range investigated, (0.5-8 kHz), compared to previous studies,
- new findings supporting theories about the OAE generation mechanism,
- the comparison between ABR and ASSR latency estimates,
- new evidence for the generation site of 80-Hz ASSR,
- suggestions about the method to choose when estimating cochlear delays,
- optimization of OAE recordings using an OAE-based chirp.

# Background and theory

---

The present study investigates cochlear delays and, in this chapter, some background and theory about generating mechanism and methods to study the cochlear delay are presented. The first section introduces otoacoustic emissions (OAEs) and presents their main characteristics and the different theories about their generation mechanisms. The second section focuses on auditory evoked potentials (AEPs), especially the auditory brainstem and steady-state responses. The sections are intended as an overview of the different mechanisms and paradigms involved. The interested reader can find more details in the references, especially Robinette and Glatcke (2001) for OAE, Katz (2001) and Hall (2006) for AEP.

## 2.1 Otoacoustic emissions

Otoacoustic emissions are low-level signals emitted in the cochlea and recorded in the ear canal. This epiphenomenon of a normally functioning auditory system was hypothesized already 60 years ago by the biophysicist Thomas Gold (Gold, 1948). His proposal of an active cochlear process was not widely accepted until the 1970s, when further evidence was found that supported Gold's theory: squirrel monkey cochleae were found to produce unexpectedly high responses to low-level sounds, and the ears of members of a human family emitted audible whistles between 7 and 12 kHz although they did not suffer from tinnitus (reviewed in Kemp (2003)). The first human OAEs were recorded and tentatively explained by Kemp (1978), who used a low-level click as a stimulus and recorded the cochlear echoes in a sealed ear canal as shown in

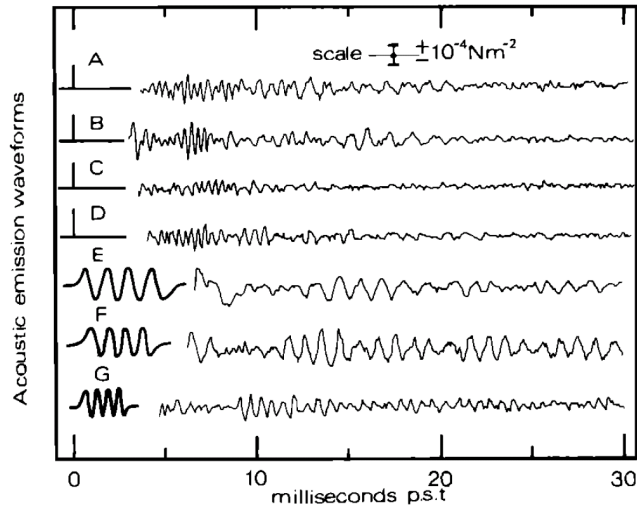


Figure 2.1: Responses recorded by D.T. Kemp in 1978, Fig. 2 from Kemp (1978). The letters correspond to different stimuli used followed by their response. Stimuli A-D are the same click sent to different ears. The OAE can be seen emerging at a post-stimulus time (p.s.t.) of around 6 ms. Stimuli E-G are 4-cycle tone bursts of 800, 1100 and 1800 Hz in the same ear. The responses in these cases resemble the stimulus, and their latency decreases as the stimulus frequency increases. Reprinted from J. Acoust. Soc. Am., 64(5), 1386-1391 with permission. ©1978, American Institute of Physics.

figure 2.1. This plot shows the amplitude of the responses to different stimuli in the time domain. Stimuli A-D are the same click sent in different ears. The OAE, in each case, can be seen emerging at around 6 ms after the stimulus onset. Stimuli E-G are 4-cycle tone bursts of 800, 1100 and 1800 Hz in the same ear. The responses in these cases resemble the stimulus, and their latency decreases as the stimulus frequency increases. The recorded sounds must have travelled through the cochlea and been re-emitted since their onset latency (ca. 6 ms for stimuli A-D) is too long to be caused by reflections in the outer or middle ear. Another observation Kemp made was that there was a non-proportional relationship between stimulus level and response level. When the input level was decreased, the response level also decreased, but it did so nonlinearly. This behaviour is an evidence of the compressive nonlinearity of the OAEs and it shed light on the link between OAEs and the amplification mechanism of the cochlea. Since Kemp's experiment, OAEs have been extensively studied and



physiological observations such as loudness enhancement and noise-induced hearing losses could then be explained.

Classically, OAEs have been classified in terms of the stimulus with which they are evoked, i. e. tones, tonal complexes, transients or no stimulus. Spontaneous OAEs (SOAEs) occur in about 40 % of normal ears when no evoking stimulus is presented (Patuzzi, 1996). To date, it is not entirely known what causes SOAEs. They may be a consequence of irregularities in the OHC-distribution pattern (Probst et al., 1991). Since SOAEs are a continuous phenomenon, no time information can be extracted from their study.

If the evoking stimulus is a low-level tone or a tone slowly sweeping a frequency range, the resulting emission is called a stimulus frequency OAE (SFOAE). SFOAEs have been measured in humans (Kemp and Chum, 1980; Brass and Kemp, 1991; Shera and Guinan, 1999, 2003; Konrad-Martin and Keefe, 2003; Schairer et al., 2006) as well as in other mammals (Shera and Guinan, 2003; Siegel et al., 2005). It is possible to estimate SFOAE group delay,  $\tau_{SFOAE}$ , by using a phase-gradient method<sup>1</sup>. This requires rather complicated post-processing.

If a stimulus is made of two tones at frequencies  $f_1$  and  $f_2$ , a typical response contains frequencies that are not present in the evoking stimulus. Due to the nonlinear properties of the cochlea, a series of distortion products of the two frequencies  $f_1$  and  $f_2$  are elicited, among which  $2f_1 - f_2$  is the largest. This type of emission is called a distortion product OAE (DPOAE). DPOAEs have been measured in various studies investigating the cochlear delay (Whitehead et al., 1996; Hoth and Weber, 2001; Schoonhoven et al., 2001; Purcell et al., 2003; Konrad-Martin and Keefe, 2003; Purcell et al., 2006). DPOAE latencies are normally estimated with a phase-gradient method and are usually measured at a high stimulus level (70-80 dB SPL).

A third category of stimulation includes transient signals like clicks or tone bursts (TB), producing so-called transient evoked OAEs (TEOAEs). Examples of TEOAEs are shown in Fig. 2.1 where the stimuli A-D are clicks and stimuli E-G are tone bursts. For the 800 Hz OAE tone burst (E), the peak occurs around 15 ms. For the 1800 Hz OAE tone burst (G), the peak occurs after approximately 10 ms. This characteristic

---

<sup>1</sup>  $\tau_{SFOAE} = -\frac{\delta\phi(\omega)}{\delta\omega}$  where  $\phi(\omega)$  is the phase of the SFOAE and  $\omega$  the angular frequency.

of the TEOAE, in which the response peak occurs at shorter time delays for higher frequencies, is called the frequency dispersion. This characteristic makes it possible to measure the latency of stimuli with different frequency components. The latency is found by visually detecting the OAE in the recorded signal. This is an easier method than deriving the latency from phase-gradient as it is the case with the group delay. Previous studies have investigated the latencies of TEOAEs in time (Norton and Neely, 1987; Şerbetçioğlu and Parker, 1999; Kapadia and Lutman, 2000; Hoth and Weber, 2001; Goodman et al., 2004; Thornton et al., 2006) and in the time-frequency domain (Elberling et al., 1985; Probst et al., 1986; Tognola et al., 1997; Lucertini et al., 2002; Sisto and Moleti, 2002; Jędrzejczak et al., 2005; Moleti et al., 2005; Sisto and Moleti, 2007). A detailed description of S-, SF-, DP- and TEOAEs is reviewed in Probst et al. (1991).

One of the main challenges of OAE research is a detailed understanding of their generation mechanism. Extensive research has been done trying to model the mechanics of the cochlea but the precise generation site(s) of OAEs, and the way they travel back to the ear canal, still remain partly unknown.

### **2.1.1 Generation mechanisms**

The mechanisms involved in the generation of OAEs differ depending on the type of OAE considered. Since only TEOAEs will be used in this work, only their generation mechanism will be addressed here. SFOAEs are thought to be generated similarly as TEOAEs whereas DPOAE generation is supposedly very different from TEOAEs generation (Probst et al., 1991; Shera and Guinan, 1999).

The most accepted hypothesis of the generation of TEOAEs is the reflection theory, suggested by Kemp (1978). It suggests that OAEs arise from reflections of the forward travelling wave (TW) in the cochlea. These reflections can occur passively and actively. A passive reflection is a simple echo, the vibrations of the stapes produce a travelling wave in the cochlea which is reflected from the cochlear partition, just like reflections from walls in a room. An active reflection refers to a reflection of the TW during its amplification by the cochlear amplifier, creating a backward travelling wave that makes the middle ear vibrate. This vibration produces a sound wave which can be recorded in the ear canal. Although this latter reflection mechanism is widely

accepted, there remain different hypotheses to explain where this reflection takes place. It has been surmised that reflections can occur: 1) at isolated points along the BM (Zwicker and Lumer, 1985), 2) at reflection sites periodically located along the BM (Strube, 1989) or 3) at reflections sites randomly spaced on the BM (Zweig and Shera, 1995). These three hypotheses are encompassed into the so-called "place-fixed" theory, suggested by Kemp (1986). This hypothesis suggests that the perturbations leading to OAEs are already in the cochlea and do not depend on the incoming wave (as the "wave-fixed" theory suggests). Following the "place-fixed" theory, the hypothesis suggested by Zweig and Shera led to the coherent reflection filtering (CRF) theory which states that, at low levels, OAEs originate from coherent reflection due to random impedance perturbations mostly from the sites of maximum excitation of the BM (Shera and Guinan, 1999; Kalluri and Shera, 2007). Depending on the reflection site location, the latency of the generated OAE will vary. For instance, the CRF theory suggests that the OAE delay is twice the delay between the stimulus onset and the peak of the (forward) travelling wave, just like a round-trip. This theory will be discussed and challenged further on in this thesis, in the light of new experimental results.

### **2.1.2 Effect of external factors on OAEs**

An interesting characteristic of TEOAEs is that their amplitude depends on the stimulus level. The input/output function of the OAEs shows a saturation when the stimulus level is increased above about 50 dB peSPL (see Stover and Norton, 1993, Fig. 4). This nonlinear growth of TEOAEs results from the amplification mechanism of the BM, which leads to high amplification at low input levels and low amplification at high input levels (see Kemp, 1978; Probst et al., 1991). Because of this mechanism, humans are able to hear very faint sounds. The nonlinear characteristic of the TEOAE is often used to differentiate OAE from noise, as it will be presented in the next chapter. The stimulus level also has an effect on the latency of OAEs. Studies have shown that an increase of the OAE level leads to shorter OAE latencies (Kemp, 1978; Wilson, 1980a; Norton and Neely, 1987). Some studies have investigated the effect of ototoxic drugs known to destroy OHCs. The intake of these drugs leads to a reduction or even suppression of OAEs, proving that OAEs are linked with the active mechanism of the cochlea (Probst et al., 1991). For the same reasons, OAEs cannot be recorded for

frequencies where the hearing threshold is worse than 30 dB HL at the respective frequency (Probst et al., 1991).

### 2.1.3 Clinical usage of OAEs

A usual recording procedure for OAEs is shown in figure 2.2. The stimuli can be generated by a computer, then sent to a D/A converter. The analog signal is then transmitted to the ear canal via an insert earphone. A microphone is placed in the sealed ear canal and connected to an amplifier. The recorded signal is then stored on a computer via an A/D converter.

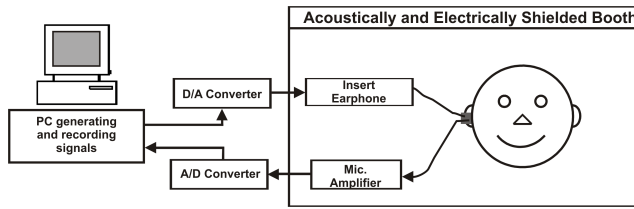


Figure 2.2: Equipment used for the recording of OAEs. The stimuli can be generated by a computer, then sent to a D/A converter. The analog signal is then transmitted to an insert earphone. A microphone is placed in the sealed ear canal and connected to an amplifier. The signal is then recorded on a computer via an A/D converter.

As mentioned before, OAEs are linked with the active mechanism of the cochlea. Their relation with the hearing threshold qualifies them for clinical tests. Click-evoked OAE are nowadays routinely used to detect hearing disorders (Kemp et al., 1986; Probst et al., 1991; Robinette and Glatke, 2001). OAEs offer a non-invasive method to measure cochlear activity without active participation of the subject. This makes OAEs an objective technique that is of importance for clinical tests with children and newborns.

## 2.2 Auditory evoked potentials

Auditory evoked potentials (AEPs) can be defined as signals originating from neurons along the auditory pathway as a response to a specific stimulus. As described in the

Responses	Abbreviation	Anatomic source	Latency [ms]
Cochlear microphonic	CM	Hair cells	<1
Summating potential	SP	Hair cells	<1
Action potential	AP	Auditory nerve	1
Auditory brainstem response	ABR		
Wave I,II		Auditory nerve	1-2
Wave III-VII		Brainstem	2-12
Frequency following response	FFR	Brainstem	6
Auditory steady-state response	ASSR	Brainstem	Stim. dep.
Middle latency response	MLR	Thalamus, auditory cortex	10-60
Late AEP	LAEP	Cerebral cortex	80-250
Mismatch negativity	MMN	Cerebral cortex	150-275

Table 2.1: Classification of the auditory evoked responses according to their latency. The first potentials to be recorded are of cochlear origin (CM) and the later responses are generated at higher levels in the brain (cerebral cortex). ASSRs are mentioned, although they can be generated at different places, depending on the stimulus characteristics. Here, the brainstem is taken as an example of an ASSR generator.

previous chapter, the auditory pathway consists of the outer, middle and inner ear, and also of the auditory brainstem and the forebrain (thalamus, auditory cortex). AEPs are recorded by placing electrodes on the scalp. These surface electrodes record the electrical activity conveyed by the skin and coming from neural generators located deeply inside the head, their signal propagating out to the scalp. AEPs are therefore a summation of electrical signals stemming from different stages along the auditory pathway and also potentials from other part of the brain. Waveforms coming from a peripheral stage of the auditory pathway will be recorded first and those coming from higher stages will be recorded later. It is therefore partly possible to separate the different components in time. AEPs have short latencies, on the order of milliseconds. Table 2.1 shows the different types of AEPs, their origins and their typical latency.

The first responses to be recorded originate from the cochlea and are recorded by electrocochleography (ECochG). The recording electrode is placed either in the ear canal close to the tympanic membrane or on the promontory of the cochlea, near the oval window. This signal generally appears 2 to 3 ms after the stimulus (usually a click). The first component is called the cochlear microphonic (CM) and is believed to be generated by the cochlear hair cells. It is followed by the summating potential (SP) and action potential (AP). Unlike the SP which is generated by the inner hair cells of the cochlea, the AP stems from the auditory nerve. The AP is also seen in

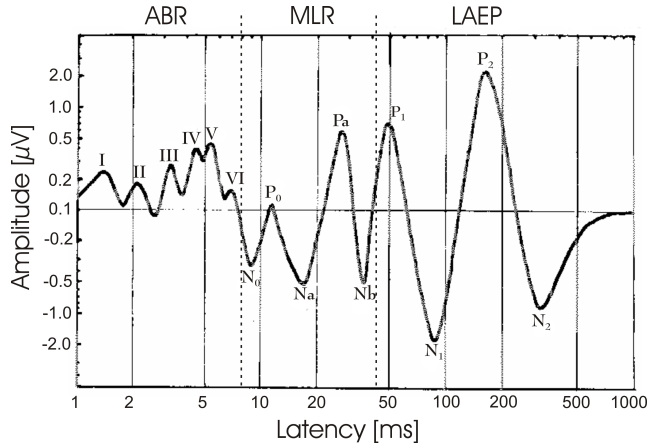


Figure 2.3: Main auditory evoked responses. The ABRs (waves I-VI) are among the first to emerge with wave V as the most dominant feature. ABRs are followed by middle latency responses (MLRs), which are a series of larger waves ( $N_0$ ,  $P_0$ ,  $N_a$ ,  $P_a$ ,  $N_b$ ). The late AEPs ( $P_1$ ,  $N_1$ ,  $P_2$ ,  $N_2$ ) present even larger waves due to the proximity of their generator (auditory cortex) with the recording electrodes. Figure modified and reprinted from Picton et al. (1974), ©1974, with permission from Elsevier.

the later auditory brainstem response (ABR) as a peak known as wave I. A typical ABR is generated in the brainstem and is composed of five to seven characteristic waves as shown in figure 2.3. ABRs have a latency between 1 and 12 ms, measured as the delay after the stimulus. A more detailed description of the ABR will be given in section 2.2.1 as it is of relevance to this thesis.

The ABR is followed by the middle latency response (MLR) which has a latency between 12 and 50 ms. The MLR is assumed to be generated in the thalamus and auditory cortex, which explains why the MLR is affected by the subject's conscious state. The late AEP (LAEP) appears in a time window of about 50 – 500 ms and is generated in the auditory cortex. Auditory steady-state response (ASSR) reflects another type of recorded responses using stationary stimuli. It will be described more in details in a later section.

Figure 2.4 shows a typical apparatus used to record auditory evoked responses. The electrodes placed on the head record all electrical signals received at the surface of the head. Brain responses are signals of very small amplitudes ( $\sim 1 \mu V$ ) and are embedded in background noise generated by spontaneous neural activity in the brain.

It is therefore necessary to average and amplify the recordings to detect the specific waveforms. The averaging is based on the assumption that: 1) the background noise is random whereas the responses due to the auditory brain activity are time-locked to the stimulus and have a constant amplitude, 2) the noise and the AEP are independent from one another. The signal-to-noise ratio (SNR) hence increases by averaging the responses. On average, the SNR is proportional to the square root of the number of averages, assuming that the noise is a stationary signal.

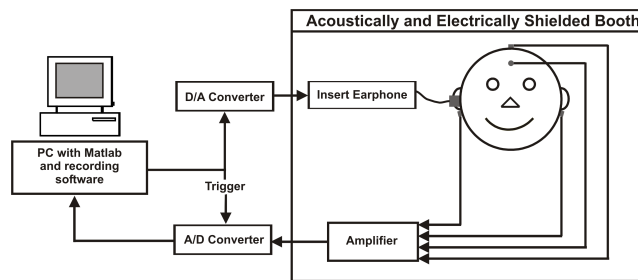


Figure 2.4: Equipment used for the recording of AEPs. The stimuli are produced by a computer, then sent to a D/A converter. The analog signal is transmitted to an insert earphone. The responses are recorded by an amplifier and accompanying software. The electrodes can be placed at different positions on the subject. The example illustrated here is to record ABRs.

Many aspects affect the quality of the recordings. On the subject side, the age, gender, state of arousal, body temperature, the potential use of drugs and the pathology of the auditory system have an influence. These factors affect each AEP differently and will be considered in sections 2.2.1 and 2.2.2. On the equipment side, the recording booth where the subject is placed and the electronic devices play a role in the recording of the signals. The noise floor should be as low as possible to help detect the characteristic waves. The electrodes need to have a good connectivity with the skin (i.e. a low impedance). The earphone with which the stimulus is presented to the ear should have a flat frequency response. Finally, the duration, intensity, polarity, frequency content and rate of the stimulus have an effect on the AEP. The optimal stimulus parameters are dictated by the kind of AEP to be recorded, e.g. LAEP needs to be elicited at a very slow rate (1 per second or less) whereas ABR can be elicited at 23/s. The optimal parameters to use will be presented in the following sections.

### 2.2.1 Auditory brainstem responses

Although the recording of electrical signals from the brain dates back to the first half of the twentieth century (Berger, 1929), the auditory brainstem responses were first described in humans in 1970 by Jewett et al.. The ABR can sometimes be referred to as a brainstem auditory evoked potential or response (BAEP or BAER) or also as a brainstem evoked response (BSER). The acronym ABR will be used in this thesis to be in agreement with the majority of the publications. A key discovery made by Jewett and Williston (1971) was to identify the main characteristics of the ABR while also suggesting the now widely employed terminology using roman numbers.

ABR, like other AEP, reflects the synchronous activation of neurons in the auditory pathway. It has long been believed that each wave of the ABR corresponds to the activity of one site, but the reality is more complex. Different sites can be responsible for a single wave or a single site can generate many waves. It seems that the earlier components originate from well localized areas, whereas it is not so clear for later waves. Wave I stems from the distal portion of the afferent cochleo-vestibular nerve (VIII<sup>th</sup> nerve), shown by Jewett and Williston (1971) and Møller and Jannetta (1983). Wave II is generated by the proximal end of the VIII<sup>th</sup> nerve as it connects to the brainstem (Møller and Jannetta, 1983), see also figure 2.5 for an illustration. From wave III on, the generation site becomes less obvious. It is believed that wave III arises from an area near or inside the cochlear nucleus, which is a bundle of neurons situated at the bottom of the brainstem. Wave IV is often merged with the more prominent wave V and has not been studied very thoroughly. It is generally accepted that wave IV is mostly due to activity in the superior olivary complex. Wave V is the wave with the highest amplitude and therefore the most easily detectable. Wave V is attributed to neuronal activities between the lateral lemniscus and the inferior colliculus on the opposite side of the stimulus (Hall, 2006). The generators of waves VI and VII are currently not clearly defined. It has been hypothesized that these waves could be the result of activity in the thalamus and also in the inferior colliculus.

Don et al. (1998) have defined the ABR peak latencies as the sum of different delays which are represented in figure 2.5:



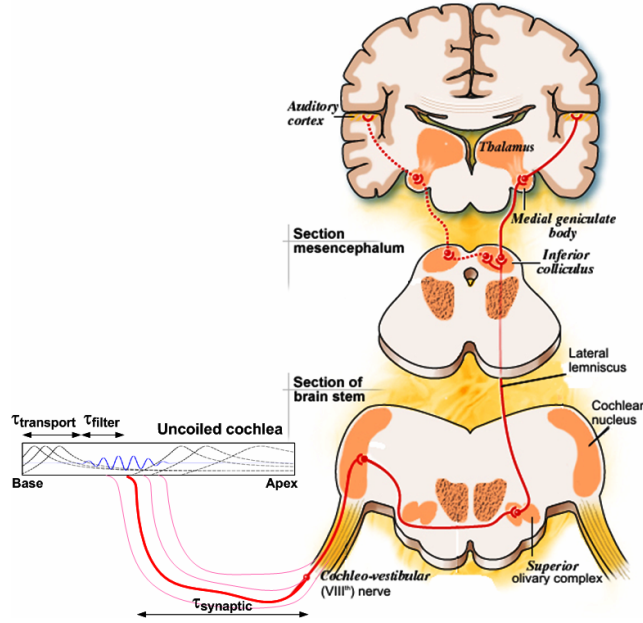


Figure 2.5: Human auditory pathway. The vibration from the basilar membrane is transmitted to the brainstem via the cochleo-vestibular nerve (also known as the VIII<sup>th</sup> nerve). The different waves observed in an ABR recording are believed to be generated at different sites. For example wave III near the cochlear nucleus and wave V between the lateral lemniscus and the inferior colliculus. The latency of the different waves is defined as the sum of the transport time ( $\tau_{transport}$  due to the passive mechanism of the cochlea), the filter build-up time ( $\tau_{filter}$  influenced by the cochlear filters), the synaptic delay ( $\tau_{synaptic}$ ) and the neural delay (not shown). Modified from a drawing by S. Blatrix, available at [www.cochlea.org](http://www.cochlea.org), used with the permission of the author.

1. The cochlear transport time,  $\tau_{transport}$ , which is due to the passive reaction of the BM, a purely linear mechanism. Other authors have denoted it  $\tau_{BM front}$  or  $\tau_{front}$ , this latter will be used in the next chapters.
2. The filter build-up time,  $\tau_{filter}$ , influenced by the sharpness of the cochlear filters, due to the cochlear amplifier. Together with  $\tau_{front}$ , they define the cochlear delay,  $\tau_{BM}$ .
3. The synaptic delay,  $\tau_{synaptic}$ , between the hair-cells activity and the auditory nerve fibres firing.

4. The neural conduction time,  $\tau_{neural}$ , between the auditory nerve activity and the place generating the ABR wave.

Unlike many clinical tools used in audiology, ABR recordings do not require the active participation of the patient. That is why they are preferred for hearing diagnosis for cases where the subject is unable to respond. These subjects include newborns, young infants and disabled or unconscious patients. ABR measurements can be used for threshold estimate or otoneurological diagnosis (e. g. VIII<sup>th</sup> nerve tumor, auditory brainstem dysfunction).

Different types of stimuli can be used to elicit an ABR. For clinical purposes, a 0.1 ms click is commonly used. The very short duration of the click provides a very broad spectrum, thus exciting the entire BM (Don and Eggermont, 1978). However, the response to a click is mainly due to nerve firing in the high frequencies region (Don and Kwong, 2002), and it is therefore impossible to extract information from lower frequencies regions of the BM. This is due to the higher travelling wave velocity at high frequencies which induces a stronger neural synchrony whereas for low frequencies, the neural activity is cancelled by activity from more basal regions.

One type of stimulus is a chirp, which takes into account the cochlear dispersion. A chirp spans a frequency range from low to high so that, by compensating for the BM travelling time, it simultaneously excites the whole cochlea. Compared to a click, this has the advantage of producing an ABR with higher amplitude (Dau et al., 2000).

Tone bursts are more frequency specific than clicks and broadband chirp and activate narrower regions of the cochlea, producing neuronal synchronicity in a restricted frequency region. However, for very low frequencies (below 250 Hz), this synchronous nerve firing is difficult to achieve since the frequencies are spread over a larger area at the apical end of the cochlea. The use of tone bursts for research purposes dates back to the early 1970s (Jewett and Williston, 1971; Terkildsen et al., 1973).

Another technique is the masking technique whose principle is to simultaneously present a tone or a click with a masker (high-pass, low-pass filtered or notched noise). The noise masker excludes the contribution from a specific region of the cochlea to the ABR. This technique was introduced by Teas et al. (1962) in animals and has since often been used in humans (Hecox et al., 1976; Davis and Hirsh, 1976; Don and Eggermont, 1978; Eggermont, 1979b; Eggermont and Don, 1980; Donaldson and

Ruth, 1993; Schoonhoven et al., 2001). This technique is not so straightforward since it requires the simultaneous presentation of a high-pass masking noise and a tone burst or click. Due to the waveform subtraction applied, it can lead to a low signal-to-noise ratio (Murray et al., 1998). It has also been found that a tone burst masked by high-pass noise brings a noticeable delay in ABR (Kileny, 1981; Hecox and Deegan, 1983; Beattie and Boyd, 1985). This is an important drawback in an experiment aiming at estimating ABR delays.

Another parameter influencing ABRs is the stimulus repetition rate. ABRs are generally recorded at a repetition rate around 20/s. The latency of the wave V increases modestly for higher rates (50 or 90/s) (Hall, 2006). The stimulus level also affects wave-V latency. It decreases when the stimulus level increases (Gorga et al., 1988).

ABRs are also listener dependent. A decrease of the body temperature produces a shorter wave-V latency and a lower peak amplitude (Hall et al., 1988). Females have been shown to have shorter response latencies and greater peaks than males (Don et al., 1993; Burkard and Secor, 2002). Age also plays a role in ABRs. Infants have longer latencies compared to adults (Katz, 2001). ABRs can also be affected by a number of pharmaceutical drugs and narcotics (Squires et al., 1978; Dixit et al., 2006) but remain independent of the state of arousal of the subject (Picton et al., 1974). In the case of hearing impairment, ABRs will be affected according to the type of impairment and the stage at which it occurs. For example a high-frequency hearing loss will lead to a delayed wave V peak. Another example is a VIII<sup>th</sup> nerve tumor which often leads to a lower ABR amplitude (Katz, 2001).

### 2.2.2 Auditory steady-state responses

In the early 1980s, Galambos et al. (1981) investigated responses to a 40 Hz modulated tone. The result was "a single, stable, composite wave", thought to be the summation of middle latency responses. Since this study, different groups explored responses evoked with tones modulated by different frequencies. The diversity of the research in this field led to various designations of the auditory steady-state response (ASSR). They can be found in the literature as envelope following responses (EFR, Dolphin and Mountain, 1992), steady-state evoked potentials (SSAEP) or steady-state evoked

responses (SSER, Rickards et al., 1994) and amplitude-modulation following response (AMFR, Kuwada et al., 2002). Throughout this thesis, these types of responses will be referred to as ASSRs. The primary purpose of the ASSR is to be an alternative technique to ABR. ASSRs can be seen as tone-burst ABRs where the stimuli are closer to each other. In other words, the repetition rate is higher. For instance, in a 205 ms long epoch consisting of tone bursts repeated at 88 Hz (one every 11.4 ms), each tone burst will evoke an auditory response similar to the one shown in Fig. 2.3. The tone-burst train will hence evoke a series of these responses. If the epoch is repeated a thousand times and averaged, the result will be a series of peaks occurring every 11.4 ms, i.e. with a frequency of 88 Hz. The frequency spectrum of the averaged response should then show a peak at the modulation frequency  $f_m = 88$  Hz, in the case where the tone-burst train has elicited a response in the brain. An example is illustrated in figure 2.6 where the stimulus is shown in the top panel and the spectrum of the response in the bottom panel. The carrier frequency of the stimulus only affects the latency of the response and not its frequency. In this case, a 4 kHz tone is modulated at a rate of 88 Hz.

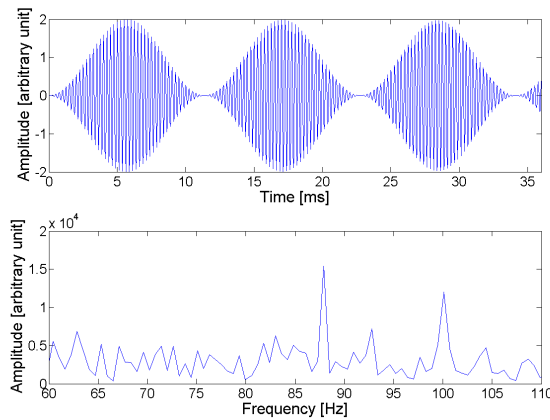


Figure 2.6: [Top] Example of a stimulus used to evoke an auditory steady-state response in the time domain. This case presents a 4 kHz tone, fully modulated in amplitude, at a rate of 88 Hz. [Bottom] ASSR measured in a normal-hearing adult displayed in the frequency domain. A peak appears at the modulation frequency ( $f_m = 88$  Hz) and also at 100 Hz, the latter being a recording artifact (second harmonic of line noise).

An analysis of the ASSR in the time domain is inconvenient since it presents a summation of waveforms and no information can be directly extracted from these time series. For this reason, ASSR measurements are considered in the frequency domain. The strength of the peak at the modulation frequency is dependent on parameters such as the subject's hearing threshold and the environment in which the measurement is made. This will be discussed in the next section.

Like other responses, ASSR recordings contain noise. Therefore, it is necessary to verify that the peak at  $f_m$  is due to the stimulus and not to the background electrical activity. One method combines the measure of the amplitude and the phase of the spectrum to detect the response and assess its robustness. The robustness is defined as the strength of the ASSR relative to the background neural noise. Studies have indeed shown that combining spectral and phase analysis enhances the detectability of ASSRs (Dobie and Wilson, 1989; Picton et al., 2001; Sininger and Cone-Wesson, 2002). A technique was introduced by Dobie and Wilson (1989) that involves calculating the magnitude-squared coherence (MSC). If  $x(t)$  is the recorded signal and  $X(f)$  its Fourier transform, the projection onto the unit circle of  $X(f)$  gives  $\cos \phi$  and  $\sin \phi$  where  $\phi$  is the phase of the signal. When  $x(t)$  is divided into  $q$  subaverages, the MSC is given by:

$$MSC = \frac{(\frac{1}{q} \sum_{i=1}^q A_i \cos \phi_i)^2 + (\frac{1}{q} \sum_{i=1}^q A_i \sin \phi_i)^2}{\frac{1}{q} \sum_{i=1}^q A_i^2} \quad (2.1)$$

where  $A_i$  is the amplitude of the  $i^{th}$  element, one of the  $q$  subaverages.

When the response is random, the phase of each average is different, and the numerator becomes close to zero. The coherence is thus low, as it would be if the signal was noise. In contrast, if all responses are identical at frequency  $f_m$ , the ratio becomes 1. The MSC method will be used to analyse the ASSR measurements described in chapter 6 and its implementation will be further discussed at that point. This method shows the importance of phase and amplitude to assess the robustness of the ASSR. These two parameters are influenced by the subject and the stimulus characteristics. A discussion of this is presented in the next section.

Studies have shown that the ASSR is generated in different brain regions, depending on the modulation frequency of the stimulus (Sininger and Cone-Wesson, 2002;

Kuwada et al., 2002). Neurons phase-lock to the envelope of an AM tone (Schoonhoven et al., 2003) and for modulation frequencies below about 20 Hz, mostly neurons from the primary auditory cortex react. For modulation frequencies between 20 and 50 Hz, the auditory midbrain, the thalamus and the primary auditory cortex are mainly involved in the response; these are also the sites where middle latency AEPs are generated. Finally, for modulation frequencies above about 50 Hz, the brainstem is mainly responsible for the generated response. Experiments have been reported that show activity up to about 260 Hz in humans (Kuwada et al., 2002).

As for other AEPs, the latency of ASSRs is shorter for higher stimulus centre frequencies (CF), due to the dispersive processing in the cochlea. Also, higher stimulus levels lead to shorter ASSR delays (Cohen et al., 1991; John and Picton, 2000), due to increased neuronal synchrony.

Cohen et al. (1991) found that the effect of sleep on ASSR depends on the modulation frequency. The decrease in ASSR amplitude due to sleep is less pronounced for  $f_m$  above 80 Hz. That same study also showed that sleep has a very small effect on ASSR latency. During sleep, the neuronal background noise tends to decrease, leading to higher SNRs. Levi et al. (1993) studied the effect of age on ASSR robustness and found that ASSRs from one-month old infants are less robust than in adults. They also concluded that ASSRs to signals modulated at 40 Hz are more robust than at higher modulation rates, reflecting a larger amplitude. This was consistent with previous findings (Galambos et al., 1981; Cohen et al., 1991). It is possible to record ASSRs from AM tones presented simultaneously. Lins and Picton (1995) showed that there is only a small decrease in amplitude of the ASSR when up to four stimuli are presented simultaneously in one ear. But the overall recording time is reduced.

### 2.2.3 Comparison between ASSRs and ABRs

The stimuli used to record ASSRs have been AM tones which excite a narrow region of the cochlea. The "equivalent" stimuli in the ABR paradigm are tone bursts. Cone-Wesson et al. (2002) and van der Reijden et al. (2006) compared the threshold estimates obtained with these two techniques. The latter conclude that "to achieve objective, frequency specific measurements, ASSR is the better choice in awake adults" whereas Cone-Wesson et al. (2002) found that the advantage of ASSR over

ABR for threshold determination is less obvious. However, a few assets of ASSR can be highlighted:

1. The peak obtained in the frequency domain for an ASSR is usually sharper and higher than the wave V from TBABR in the time domain. Therefore, it is easier to detect. Besides, a robustness method such as MSC renders ASSR more advantageous over ABR, for which fewer methods exist to detect the waves.
2. ASSR stimuli are more frequency specific than tone bursts. This is due to their duration, which produces a narrower spectrum and no second lobe excitation as is the case for TBABR.
3. ASSRs may be a bit faster to record and simultaneous recordings are possible, which decreases the total measurement time.

On the other hand, an undisputable advantage of ABR is the wide availability of the recording equipment, which explains their widespread use in clinics compared to ASSR. The present study will compare both ABR and ASSR measurement techniques in terms of their latency as a function of frequency.

## 2.3 Summary

This chapter gave a short presentation of three techniques to measure emissions or potentials generated within the human auditory system, namely otoacoustic emissions, brainstem responses and steady-state responses. The last two are generated by neuronal activity whereas OAEs reflect the activity stemming from "mechanical" BM processing. The goal of the present study is to estimate human cochlear delays based on these non-invasive techniques. Once delays are estimated, the methods can be compared in terms of reliability and robustness to background noise and the most appropriate method to estimate cochlear delays can be found. The actual latency estimates from the different experiments can also be compared to learn about the mechanisms involved in OAEs and AEPs generation. This will either confirm or cast doubt on current theories about cochlear mechanics and OAE propagation. This series of cochlear delays estimation starts in the next chapter with OAEs evoked by tone bursts.





# Individual cochlear delay estimates using otoacoustic emissions

---

*This work was presented at the XXVIII<sup>th</sup> International Congress of Audiology, 3-7 September 2006, Innsbruck, Austria.*

## **Abstract**

Methods to estimate cochlear delay in humans with OAEs have been traditionally based on phase-derived group delays. This method demonstrates large variability in cochlear delay estimates, and is derived from across subject averages. This work aims to assess the individual variability in cochlear delay by measuring tone-burst evoked otoacoustic emissions (TBOAEs) in 16 normal-hearing adults. The OAE is analysed by separating the non-linear components of cochlear origin, and the linear reflection in the time domain. The observed latencies as a function of frequency are qualitatively similar across subjects. For the individual subjects, the delay for each tone-burst frequency is reproducible. Defining OAE latency as the time between the onset of the stimulus and the peak of the first OAE burst yields results in agreement with previous studies. However, care must be taken when comparing the results of previous studies. This is due to an ambiguity in the time domain regarding the true onset point of the OAE, and hence the derived cochlear travelling wave latency. The inter-subject variability explains the discrepancy observed in other studies e. g. using different stimulus paradigms. The relatively small within-subject variability suggests that the present method is a good approach for estimating cochlear delay.

### 3.1 Introduction

The previous chapter presented different types of stimuli used to estimate OAE latencies. Among them, tone-bursts (TB) are narrow band stimuli that offer the possibility to record frequency-specific delays. Due to the tonotopic mapping of the cochlea, OAEs evoked by low frequency TB have a higher latency than OAEs evoked with high frequency TB. The present chapter presents the experiment conducted to estimate cochlear delays using TBOAE in normal-hearing adults.

When a signal is presented to the ear, reflections (echoes) occur in the ear canal and the middle ear. These components arise shortly after the stimulus onset and are linear. There are also reflections in the cochlea that appear later and are nonlinear. All these reflected signals are merged to form the signal recorded in the ear canal. One difficulty is therefore to detect the onset of the OAEs masked by the other reflections. In the present study, this OAE onset ambiguity will be solved by using two clicks at different levels. After normalization and subtraction of the recorded signals there only remains the nonlinear part of the response, containing the OAEs. This takes advantage of the nonlinearly growing function of the TEOAE, see section 2.1. A detailed description of this method will be given in the next section, it is similar to the subtraction method used in previous studies (Keefe, 1998; Şerbetçioğlu and Parker, 1999).

Some studies have noted a remarkable standard deviation in their averages across subjects, Shera and Guinan (2003) explains this by an "intrinsic variation" of the OAE phase. In the present study, the inter-subject variability is compared with other studies and explanations are suggested. The variation of the OAE latency within the same subject (intra-subject variability) is also of interest. This aspect of OAE measurement has not, to date, been studied thoroughly. The present study therefore investigates this intra-subject variability and suggests some explanations.

Some studies have investigated the influence of noise (either background noise or from physiological origin) on DPOAE prevalence and level (Gorga et al., 1994; Nelson and Zhou, 1996) or on SFOAE level (Schairer et al., 2003). The present study investigates the effect of noise floor on TBOAE latency estimates.

This chapter presents measurements of individual cochlear delay estimated using OAEs. The new paradigm used to resolve the OAE onset ambiguity will be presented. It will also be shown that, although highly repeatable for each subject, the OAE latencies present a sizeable variation across subjects. It will be shown that the criterion retained to limit the contamination of the recordings by noise has little effect on the OAE latencies.

## **3.2 Methods**

### **3.2.1 Subjects**

The subjects participating in this experiment were 16 normal-hearing adults: 4 females, 12 males, aged between 22 and 30 years. They all had pure-tone thresholds better than 15 dB HL in the range 0.25-8 kHz and ear canals free of wax or other debris (checked by otoscopy). All subjects were paid for their participation. They were seated in a comfortable chair in a sound insulated booth compliant with the IEC 268-13 standard. They were asked to move as little as possible during the session and to swallow only in between the recordings (ca. every 3 min.). Each session lasted about 45 min. The responses were recorded during three different sessions, with at least two hours between sessions.

### **3.2.2 Stimulus generation and response measurement**

The stimuli used were clicks and tone bursts repeated at a rate of 25/s. Clicks were presented 4000 times at 56 and 66 dB peSPL. Tone bursts were presented 4000 times at 66 dB peSPL, chosen to be comparable with historical studies. A lower level would compromise the comparison with other physiological data such as ABR, whose waves are difficult to detect. At higher levels the cochlear amplifier is not as active and the OAE produced are therefore of lower amplitude (Kemp, 2002). In that case, it is hard to extract the response from the noise. Tone-burst frequencies ranging from 500 Hz to 8 kHz were used and their durations ranged between 10 and 1.25 ms (see table 3.1). These durations represent a trade-off between having an equal number of cycles for all frequencies and a relative narrow spread in their spectrum. It was tried to reach a

similar number of cycles for all frequencies so that the basilar membrane is uniformly excited for all frequencies. Some stimuli present fractions of cycles, these have no effect on the TB bandwidth due to the hanning window applied<sup>1</sup>.

Frequency kHz	Total Length	
	ms	cycles
0.5	10	5
0.75	7	5.25
1	5	5
1.5	5	7.5
2	5	10
3	3.4	10.2
4	2.5	10
6	1.7	10.2
8	1.25	10

Table 3.1: Tone bursts used, with length in ms and in number of sinewave cycles. These values are a compromise between a narrow spread in frequency and a short stimulus in time.

The choice of the stimulus was inspired by the experiments from Norton and Neely (1987) and Şerbetçioğlu and Parker (1999) and were generated following the standard IEC 60645-3 (IEC, 2007), on short duration test signals. A time and frequency representation of two stimuli is presented in figure 3.1, this shows the frequency specificity of the stimuli and the compromise made regarding their length (in ms and number of cycles). The same plots for the nine stimuli are available in appendix A.

The stimuli were generated with Matlab and sent to a RME D/A converter (ADI-8 Pro), connected to a programmable attenuator (TDT PA5) and to the headphone driver (TDT HB7), and finally delivered to the ear canal via a miniature transducer (Etymotic ER-2). The stimuli were calibrated using an ear mould simulator (B&K DB 0370) connected to a IEC 711 coupler (B&K 4157) and a B&K 2607 sound level meter. A detailed description of the calibration is presented in appendix B. The TBOAEs were recorded in the right ear in most of the subjects but in case of a blocked ear canal, the left ear was chosen. Studies have shown that the two ears of a subject present very

<sup>1</sup> Less than 18 Hz difference between the bandwidths with 7 and 7.5 cycles, for the 1.5 kHz TB.

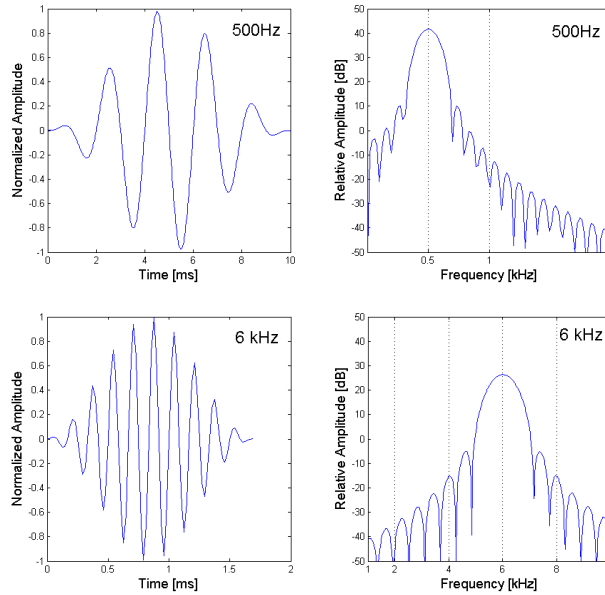


Figure 3.1: Time and frequency plot for two tone bursts, 500 Hz and 6 kHz. The spectra show how narrow the tone bursts are, a good frequency specificity insures that the neighbouring filters on the basilar membrane will not be excited.

similar TEOAE (Probst et al., 1991), the choice of the ear has therefore no influence on the recorded data. The responses from the ear canal were recorded with a ER-10B low-noise microphone (Etymotic Research), band-pass filtered 150-16000 Hz with an analog filter (Krohn-Hite 3750), digitalized with the RME A/D converter and, finally, saved on a hard drive for off-line analysis. Both stimulus generation and response measurement were controlled by a Matlab program developed in collaboration with Manfred Mauermann at the University of Oldenburg.

### 3.2.3 Off-line data analysis

#### Artifact rejection

In order to reject averages containing a great amount of noise due to subject movement or swallowing, an artifact rejection was applied. Averages contaminated by external noise can be detected due to the presence of sudden high amplitude signals, the averages were ranked according to their maximum amplitude and 10% of them were discarded.

#### Noise considerations

An important parameter of the present study is the noise floor of the recordings. One of the purposes of this study is to investigate the effect of signal-to-noise ratio (SNR) on TBOAE latency estimates. 4000 averages were recorded for each stimulus, thus it is possible to examine how the latency evolves with the number of averages. Each doubling of the number of averages ideally leads to an increase of the SNR by 3 dB (Elberling and Don, 1984). Deviations from this rule is due to the rough assumption that the noise is an ergodic random process (Beattie and Ireland, 2000). A review of the literature indicates that previous studies have mainly used two criteria to ensure a good reliability of the data: a fixed SNR (Konrad-Martin and Keefe, 2003; Shera and Guinan, 2003; Schairer et al., 2006), a number of averages to reach (Norton and Neely, 1987; Killan and Kapadia, 2006) or one of these two, which ever occurs first (Gorga et al., 1994, 2000). It has indeed been previously observed (Elberling and Don, 1984) that the SNR can be deteriorated if the number of averages is too high. Prior to the present study, the evolution of the SNR level as a function of the number of averages was investigated, see Fig. 3.2. This confirms that increasing the number of averages for a measurement does not necessarily lead to a better SNR. This trend shows that the noise is not stationary, as mentioned earlier. The signal-to-noise ratio was calculated as:

$$SNR = 20 \log \left( \frac{RMS(signal)}{RMS(noise)} \right) \text{ dB}$$

where the noise is defined as the subtraction of two subaverages divided by two and the signal is the sum of two subaverages divided by two. There is nevertheless an error to this noise calculation. The assumption made when taking sampled signals is that there is an infinite number of recordings or a recording of infinite length. This carries an inherent "statistical sampling error" given by  $(\sqrt{BT})^{-1}$  where  $B$  is the bandwidth over which the data is uniformly distributed and  $T$  the length of the recording (Bendat and Piersol, 1980). However, this error is not significant in the present study<sup>2</sup>.

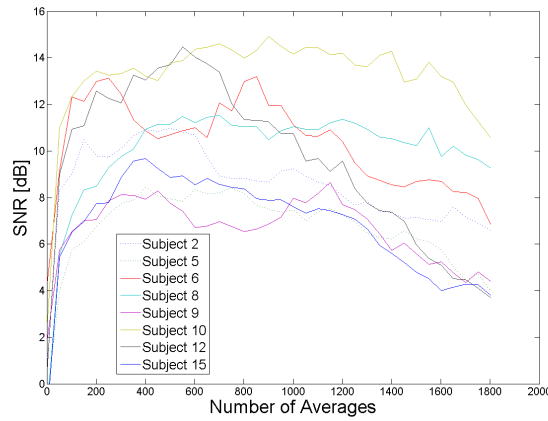


Figure 3.2: Evolution of the SNR as a function of the number of averages for several subjects at 2 kHz. This shows that increasing the number of averages for a measurement does not necessarily lead to a better SNR. A similar conclusion was also drawn by Elberling and Don (1984).

While calculating the SNR, it was noted that it was difficult to achieve the same SNR for all TB frequencies. For some frequencies it was impossible to get a SNR greater than 16 dB whereas for other frequencies (e.g. 8 kHz) 26 dB was easily reached. Gorga et al. (1994), who measured OAEs in various conditions, came to the same conclusion. This is why it was decided to take different signal-to-noise ratio for the different TB frequencies. Prior to defining the SNR limit for each TB frequencies it is necessary to see how the SNR evolve as a function of the number of averages. 4000

<sup>2</sup> For instance the error for the 1 kHz TBOAE is estimated around 0.3 dB, this is negligible compared to the OAE level (approx. 20 dB).

averages are available for each recording but, as seen earlier, not all are necessary to reach the maximum SNR. A typical OAE recording is shown in Fig. 3.3. The first part of the recording contains the recorded stimulus, the SNR is expected to be high for this part. On the contrary, the end of the signal mostly contains background noise since TBOAEs and their reflections have occurred already. This part of the recording is expected to have a low SNR. The SNR should reflect the strength or weakness of the OAE signal relative to the background noise. When there are no OAEs, outside the two dashed lines in Fig. 3.3, calculating the SNR is flawed. The first limit is known and corresponds to the end of the stimulus. Then a moving SNR is estimated and when it becomes lower than 6 dB, the second limit is set. All SNR calculations are therefore only based on the portion of signal between the two limits.

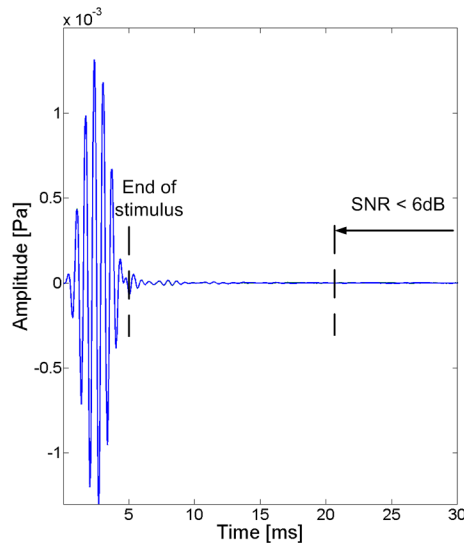


Figure 3.3: The portion of signal taken to calculate the SNR lies between the two dotted vertical lines. After the second line, the recorded signal is dominated by background noise.



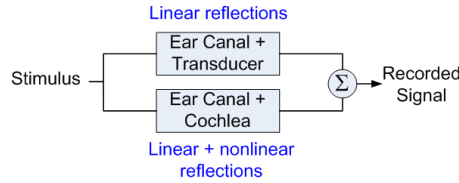


Figure 3.4: Every signal recorded in the ear canal has been through a number of stages in the auditory pathway. Each stage brings some modifications to the signal. In the case of otoacoustic emissions, the recorded signal can be considered as the sum of linear and nonlinear components.

### Separation method

The main problem encountered when analysing TEOAEs in the time domain is to separate the recorded stimulus from the OAE response. The detection of the OAE onset is hard due to early linear reflections occurring in the transducer and the ear canal (EC) (Stover and Norton, 1993). This phenomenon is sketched in Fig. 3.4, which shows that the recorded signal is the sum of different components: linear reflections of the stimulus in the ear canal, nonlinear reflections in that same ear canal and also the reflections in the cochlea which have both a linear and a nonlinear part. As explained in section 2.1, OAEs are the result of two sources: one due to passive cochlea reflection (linear reflection) and another due to hair cell based retransmission (nonlinear reflection) (Kemp, 2002).

In previous studies, the separation between linear and nonlinear reflections has been done similarly for all subjects (Kemp, 1978; Wilson, 1980a; Wit and Ritsma, 1980; Norton and Neely, 1987; Keefe, 1998; Şerbetçioğlu and Parker, 1999; Jedrzejczak et al., 2005). This assumes that all the subjects have identical ear canal and middle ear. The ear canal is also often approximated by a 2cc coupler and tests were run to see if this assumption was acceptable. The conclusion is that a 2cc coupler is only a rough approximation of the human ear canal and does not lead to a better separation of the linear and non linear components of the OAE response. There is therefore a need for an estimation of the ear canal influence on the recorded signal for each subject. The paradigm used in the present study tries to resolve the ambiguity of the OAE onset in each subject separately. This new method consists of:

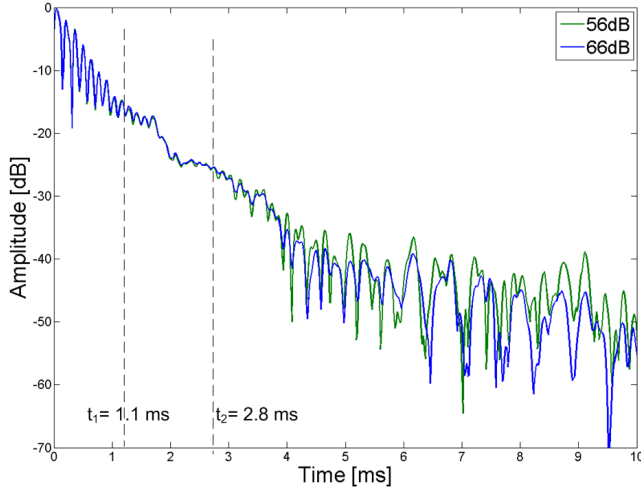


Figure 3.5: Hilbert transform of the normalized click responses at 56 dB peSPL (green line) and 66 dB peSPL (blue line). Before the first dashed vertical line ( $t_1$ ), the responses are superimposed, corresponding to a linear behaviour of the signal. After the second line ( $t_2$ ) the two curves differ, meaning that the responses contain non linear components.

1. Recordings of click-evoked OAEs at two different levels. The impulse response of the EC is obtained and this gives a filter estimate of the EC.
2. Convolution of the TB stimulus with the EC filter estimate.
3. Comparing tone-burst OAEs and the convolved signal.

Once click-evoked OAEs have been obtained at two different levels, the responses are normalized (i. e. divided by their maximum amplitude to obtain an amplitude of 1 for each recording). For two signals linearly related, this normalization brings their curve on top of each other once plotted<sup>3</sup>. On the contrary, for two signals not linearly related, their curve will be different. As shown in figure 3.4, OAE recordings are the sum of linear and nonlinear signals, normalizing these recordings helps differentiating

<sup>3</sup> if  $y = k \cdot x$  then  $y_{max} = k \cdot x_{max}$ ,  $k \in \mathbb{R}$ , normalizing  $y$  gives  $\frac{y}{y_{max}} = \frac{k \cdot x}{y_{max}}$  leading to  $\frac{y}{y_{max}} = \frac{x}{x_{max}}$ . The two normalized signals are equals.

the linear and nonlinear parts of the signal. Figure 3.5 shows the envelope<sup>4</sup> of the normalized click responses (green line: 56 dB peSPL, blue line: 66 dB peSPL) measured in one subject. The first parts of the curves are similar, meaning that the responses are linearly related, i. e. the increase in level of the stimulus leads to the same increase in level of the response. After the first dashed line (ca.  $t_1 = 1.1$  ms) the responses are not on top of each other anymore, which means that an increase in level of the stimulus leads to a different increase of the response amplitude. This is an effect of the nonlinear activity occurring in the ear canal and the cochlea. After  $t_2 = 2.8$  ms, the 56 dB curve is above the 66 dB one. This is also a characteristic of the nonlinear growth function of TEOAE (see section 2.1). This means that after  $t_2$  the response contains components of cochlear origin. The ear canal impulse response is obtained by removing echoes that are not due to the ear canal nor the transducer. That is why the signal after  $t_1$  and  $t_2$  is removed, giving two estimates of the EC impulse response. An example is plotted at the top-right corner of Fig. 3.6. Each filter estimate is then convolved with the stimulus (tone burst). An OAE response is also obtained and a least-square fit algorithm is used. This filtering technique calculates the amount of signal of a certain frequency contained in the TBOAE recording. Since OAEs have a spectrum similar to that of the evoking tone burst (Wit and Ritsma, 1980; Kemp et al., 1986), searching for the amount of TB frequency in the recorded signal will return the amplitude of the OAEs. An example can be seen at the bottom of Fig. 3.6, the first part of the signal corresponds to the recorded stimulus and has therefore a high amplitude, the following peaks and troughs correspond to the presence or absence of the sought frequency (here 750 Hz). This least-square fit algorithm is based on a study by Long and Talmadge (1997) and is further explained in appendix C.2. The algorithm is applied to the TBOAE response and the convolved signal. The two envelopes hence obtained are compared to help separating visually the OAE response from the linear reflections occurring in the ear canal (see Fig. 3.7). As figure 3.7 shows, there are a number of bursts appearing, each with different levels and latencies. The expected level of the OAEs is about 45 dB below the peak stimulus level (Wilson (1980a); Wit and Ritsma (1980) and Robinette and Glatke (2001)). This helps distinguishing the

---

<sup>4</sup> more precisely the absolute value of the corresponding analytic signal, see appendix C.1

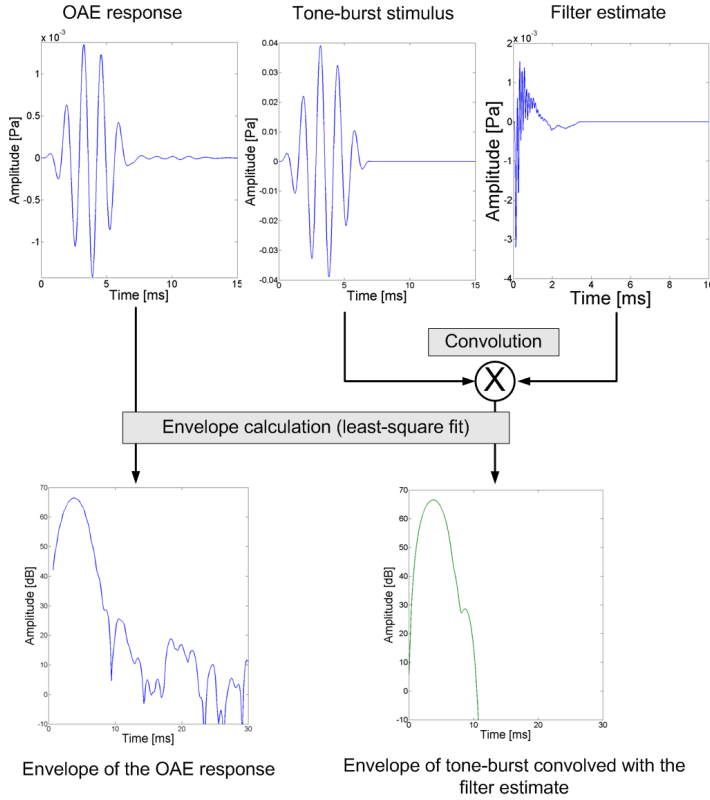


Figure 3.6: This diagram summarizes the paradigm used to help resolving the OAE onset ambiguity. The filter estimate is obtained from click responses, see Fig. 3.5. This estimate is then convolved with the tone burst to simulate the effect of the ear canal on this stimulus (i. e. the length of the linear reflections). The envelope of the resulting signal is calculated by a least-square fit algorithm, the envelope of the recorded OAE is also obtained in the same way and the comparison of the two envelopes is shown in Fig. 3.7.

OAE burst and therefore obtaining its latency. For the example illustrated in figure 3.7, the OAE latency is estimated at 10.8 ms.

### Defining the TBOAE latency

The OAE latency was defined, in the present study, as the time between the onset of the stimulus and the peak of the burst detected as an OAE, the blue peak around

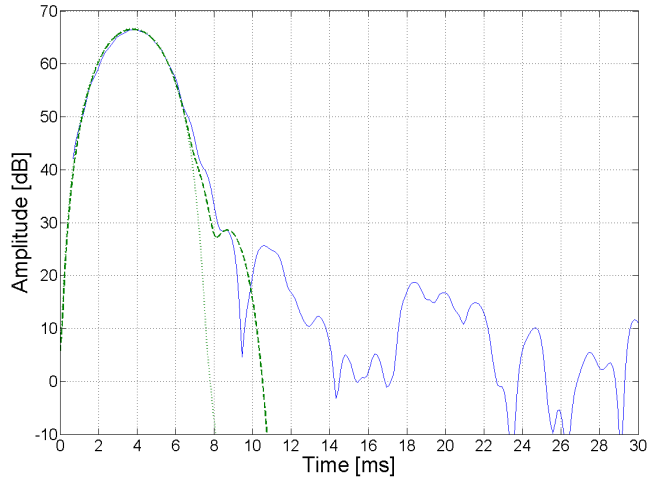


Figure 3.7: Envelopes of the convolved signal (green dotted curves) and of the OAE response (blue line). The stimulus is here a 750 Hz tone burst. As explained earlier, the amount of signal found before the dotted green line is due to linear reflections and the signal located after the dashed green line is more likely to be from cochlear origin.

10.8 ms in Fig. 3.7. With this definition, Serbetçioğlu and Parker (1999) found results close to other studies.

## 3.3 Results

### 3.3.1 Effect of frequency on TBOAE latency

A typical plot of a TBOAE envelope is shown in Fig. 3.8. The grey curve is the actual response where the beginning (containing the recording of the stimulus) has been cut out for clarity. The blue line is the envelope of the response obtained by the least-square fit method (see section 3.2.3). The two green curves represent the envelope of the estimated ear canal filter obtained with the two times  $t_1$  and  $t_2$  (see Fig. 3.5) and the red line is the phase of the response. The phase is used as a supplementary cue to detect the OAE burst. As mentioned before, reflections occur in the EC and the cochlea. The signal reflected in the ear canal and the wave originating in the cochlea have

different phase due to the different mechanisms involved in the cochlear reflection and the different lengths they have travelled (Shera and Guinan, 1999).

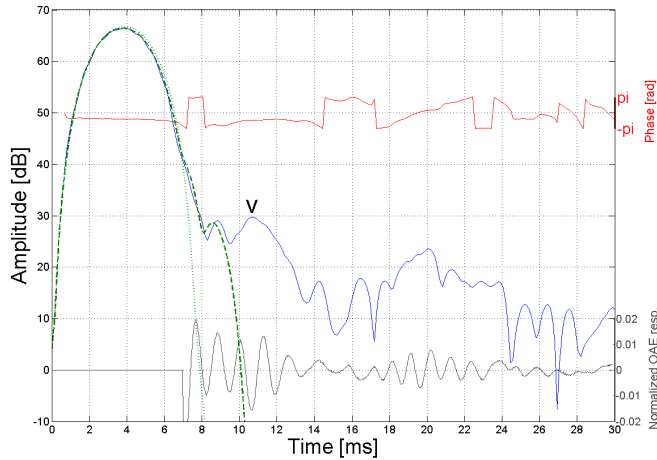


Figure 3.8: TBOAE response at 750 Hz. The blue line is the envelope of the recorded signal (gray curve) obtained by the least-square fit method. The two green curves represent the envelope of the estimated ear canal filter and the red line is the phase of the response. The gray curve represents the normalized response plotted on a different scale, it is just used to confirm the presence of emissions. The "V" marks the peak of the OAE burst.

As explained in the previous section, the signal observed before the dotted green line is mostly due to linear reflections (outer and middle ear) whereas the signal after the dashed green line presents a superposition of linear and nonlinear reflections. On the gray curve, a signal can be seen emerging around 8 ms after stimulus onset and ending at around 12.5 ms. The level of this oscillation is 40 dB below the stimulus level, its frequency is similar (see blue curve). This oscillation, or burst, is therefore identified as an OAE, the latency is taken at the maximum of this envelope (see "V" in Fig. 3.8). There appears to be a similar waveform with decreased amplitude occurring around 19 ms, this might be an echo of the OAE due to internal reflections as hypothesized in previous experiments (Kemp and Chum, 1980).

It should be noted that the method developed in the present study does not allow to detect the OAE burst with a 100% certainty. Although best efforts were made

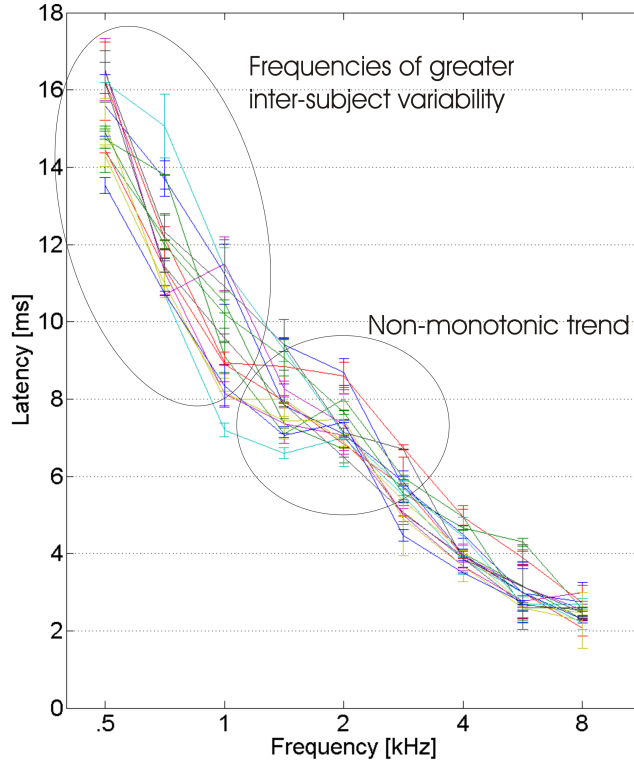


Figure 3.9: Latencies for the 16 subjects with their standard deviation  $\pm 1$  std). The two ellipses indicate area of interest. At lower frequencies the inter-subject variability is greater and around 1-2 kHz a "knee" appears in the latency of OAE. The average values for the latencies and their standard deviation are presented in appendix D.

to provide a reliable tool, the OAE onset ambiguity is not solved. This ambiguity seems to be inherent to any TEOAE recordings and is possibly not fully resolvable. The latency estimates for all subjects are shown in Fig. 3.9. The latency decreases with increasing frequency as a function of TB centre frequency, as expected from the tonotopic organization of the cochlea. Nevertheless, this trend appears non-monotonic for some subjects. Especially around 2 kHz, the latency was found to be higher than at 1.5 kHz. This has also been observed previously (see Kemp (1978); Wilson

(1980a), Johnsen and Elberling (1982) Fig. 6 and Schoonhoven et al. (2001)). This phenomenon will be further discussed in section 3.4.1.

### 3.3.2 Intra-subject variability

The repeatability of the OAE latency was verified by measuring during three different sessions. Fig. 3.9 shows the standard deviation for each subject at each frequency. The exact values for mean and standard deviation are presented in appendix D. For some subjects, the obtained values were highly repeatable (Fig. 3.10(a) for subject 7) whereas for other subjects, differences were found between the runs (Fig. 3.10(b) for subject 11). The intra-subject variability is rather small, with a maximum standard deviation of 1.2 ms (Subject 14 at 1 kHz).

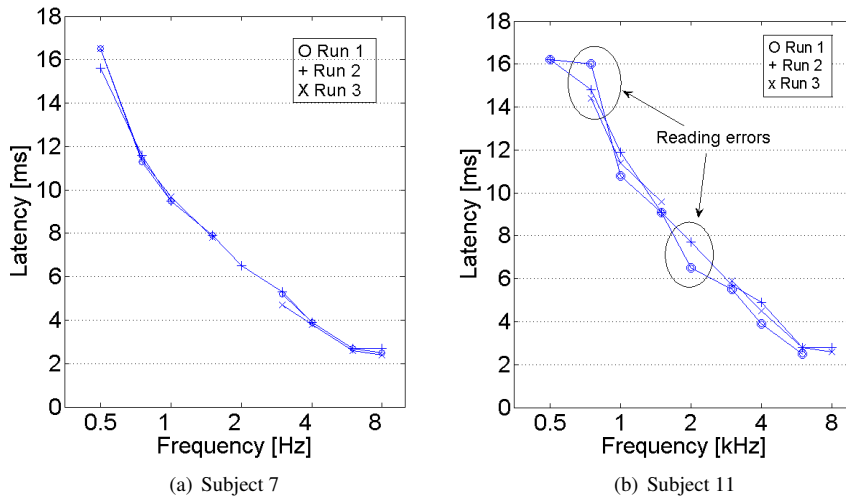


Figure 3.10: Example of two subjects whose latency repeatability differs. These are two illustrative cases. Subject 7 has a very good intra-subject variability, the three curves corresponding to the three measurement sessions are almost perfectly on top of each other. Subject 11, on the other hand, has a larger variability even though it stays relatively small. The good correlation between the runs in subject 11 shows that the discrepancy observed at 0.75 and 2 kHz is due to reading error. These two examples were obtained with a fixed SNR.



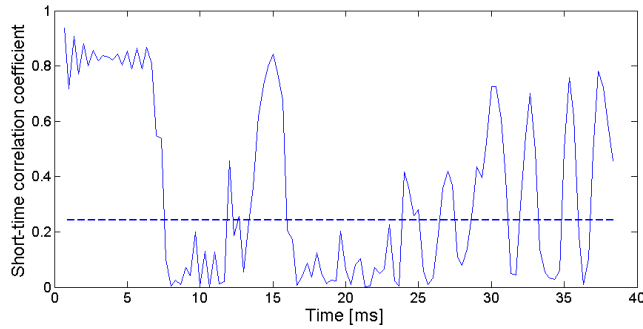


Figure 3.11: Short-Time Correlation Coefficient (STCC) between run 1 and run 3 at 750 Hz for subject 11. The dashed line represents the 95% confidence limit. If the curve is above this limit, this means that the two runs are highly similar. Therefore, a difference in the OAE latency estimate is likely due to reading errors.

A possible explanation for the poorer repeatability in some subjects could be the difficulty of assessing the correct OAE burst. Probst et al. (1991) and Norton and Neely (1987) also argued that the large range of the estimated latency at 1 kHz (10-16 ms), observed in the literature, is due to the difficulty of assessing the correct burst. It is necessary to investigate if this difference of latency between runs is due to reading errors or to actual variability of the OAEs. To verify this, a correlation method was applied. This short-time correlation coefficient (STCC) was calculated to compare the three runs at the different frequencies. If the STCC is close to 1 this means that the two signals are very similar (see Harte and Elliott (2005) and appendix C.3 for further explanation about the STCC). The STCC between run 1 (the outlier in Fig. 3.10(b)) and run 3 of subject 11 is presented in Fig. 3.11. As can be observed on this figure, the correlation is above the 95% confidence limit at the beginning (which can be expected since this corresponds to the recorded stimulus) and also around 15 ms. This proves that both runs present the same shape at that time. The consequence is that the value reported for either run 1 or run 3 is wrong. Since the value read for run 1 is clearly different from the two other runs, this value is not correct. In the same way the value found in run 3 should be taken as the TBOAE latency since it agrees with run 2. The OAE latency for this subject 11 at 0.75 kHz is therefore 14.4 ms. Calculating the STCC for the different outliers points observed for different subjects explain most of

them. This means that the observed difference (e. g. in Fig. 3.10(b)) is due to bursts mistaken for being OAE bursts.

Besides reading errors, there are a couple of other factors that can account for the intra-subject variability, such as the position of the subject's head. It has indeed been shown that the posture has an influence on the OAE strength (Johnsen and Elberling, 1982; Wilson, 1980a).

The conclusion is that the variability of the latency estimates within each subject is rather small, as verified with the STCC calculation. The worst case is a 2.2 ms difference between the three measurements. This is very small compared to the latency of the OAE at that frequency (ca. 11 ms). This intra-subject variability is more likely to be explained by the difficulty of detecting the correct OAE burst than by a change in the OAE latency.

### 3.3.3 Inter-subject variability

The difference of OAE latency estimates between subjects is an interesting feature of this experiment. As seen in Fig. 3.9, there is a noticeable variability among subjects. This is particularly true at low frequencies where the latency found for the TBOAE can differ by up to 4 ms between the individuals tested. The inter-subject variability has also been found to be quite large in previous studies (Johnsen and Elberling, 1982; Grandori, 1985; Dreisbach et al., 1998; Hoth and Weber, 2001). The standard deviation ranges, in the present study, from 7% at 2 kHz to 15% at 6 kHz. This is in agreement with values found by Norton and Neely (1987): 10% - 30% and Kemp and Chum (1980): 20%. The inter-subject variability decreases with increasing frequency, for example 2 ms at 0.5 kHz and 0.9 ms at 8 kHz. This was also observed by Hoth and Weber (2001) but no explanation was given. As seen in the previous section, the intra-subject variability obtained in this study for the TBOAE latency estimates is very small. This contrasts somehow with the rather large variability between subjects. This implies, on the one hand, that results for each individual are reliable but on the other hand there seems to be an inherent inter-subject variation. It appears therefore reasonable to analyse each subject individually rather than taking an average across them.

### 3.3.4 Effect of the noise floor on OAE latency

The latencies plotted in Fig. 3.9 were obtained with a SNR fixed for each frequency. These SNR values are the lowest that could be reached for the 16 subjects. They are summarized in table 3.2.

Frequency [kHz]	0.5	0.75	1	1.5	2	3	4	6	8
SNR [dB]	23	23	20	17	14	20	23	14	26

Table 3.2: Tone-burst frequencies and their corresponding SNR for which the latency was calculated. There does not seem to be any relations between the SNR level and the standard deviation across subjects seen in Fig. 3.9. This was verified by calculating the correlation coefficient between these two variables, the result,  $r = 0.04$ , is close to zero and confirms that there is no correlation.

There does not seem to be any relation between the SNR level and the standard deviation across subjects (correlation coefficient  $r = 0.04$ , see appendix C.4 for details about its calculation). Indeed, at 2 and 6 kHz, the SNR is lowest (14 dB) but the standard deviation is smaller than at lower frequencies. Thus, the variability between subjects is not related with the variance of the noise floor. But an advantage of taking the same SNR for all subjects at a specific frequency is to ensure that the analysed signals have an equal amount of noise. The latencies thus found have the same likelihood to be correct. In order to see the influence of the noise floor on the TBOAE latency, a comparison was made between latencies obtained at a fixed SNR (Fig. 3.9) and latencies obtained with the maximum number of averages (4000), see Fig. 3.12. The curves represent the average across subjects and the standard deviation is also shown to indicate the inter-subject variability. The blue curve is hence the average of the individual values presented in Fig. 3.9. This figure shows that the estimated delay of the TBOAE is slightly higher when using the maximum number of averages, but it stays in the range of the standard deviation of the fixed SNR (solid line). Even at 3, 4 and 6 kHz where the two curves are most separated, there is no complete distinction between the fixed SNR curve and the other one. This shows that the TBOAE delay measured with a fixed SNR is not significantly lower and both methods (fixed SNR and maximum number of averages) lead to similar results for TBOAE latency estimates. Regarding the inter-subject variability, the same trend is observed for both curves, this variability is higher at low frequencies (0.2 ms at 8 kHz vs. 0.9 ms at 0.5 kHz). However, this variability is greater for the maximum number

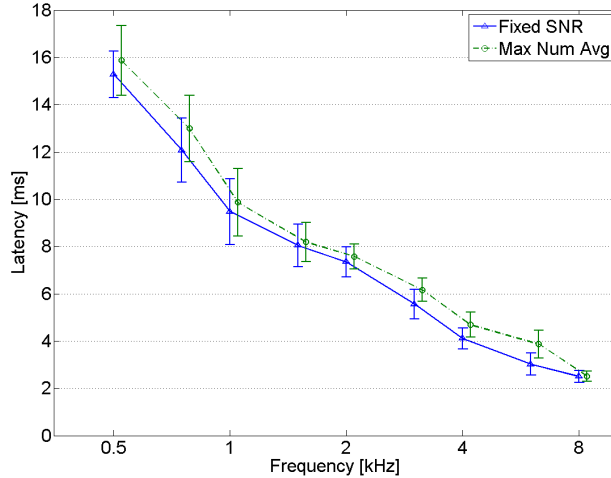


Figure 3.12: Comparison of the mean TBOAE latencies across frequencies for a fixed SNR (solid curve) and a maximum number of averages (dashed curve). Each measuring point is the mean across subjects presented with  $\pm 1$  standard deviation. The green curve has been shifted horizontally to improve visibility.

of averages (1.4 vs. 0.9 ms at 0.5 kHz); this could be attributed to breathing noise or head movements (face scratching, swallowing...) as suggested by Gorga et al. (1994). Therefore, having a different noise level for each subject increases the possibility of reading errors. Hence, for studies where the latency estimates are averaged across subjects, a fixed SNR leads to a more reliable estimate of the TBOAE delay at low frequencies.

The conclusion about the effect of the noise is that it is preferable to reach a different (but fixed) SNR for the different frequencies rather than recording the same number of averages. Since the SNR depends on the subject (strength of OAEs, quietness) it is better to estimate latencies based on equally good responses for all subjects, i. e. a different number of averages. Albeit the choice for a fixed SNR does not lead to significantly different estimates, it decreases the chance of wrong OAE bursts detection. Besides, a fixed SNR can be reached in a far less number of averages, the fixed SNR method is therefore faster than the maximum number of averages.

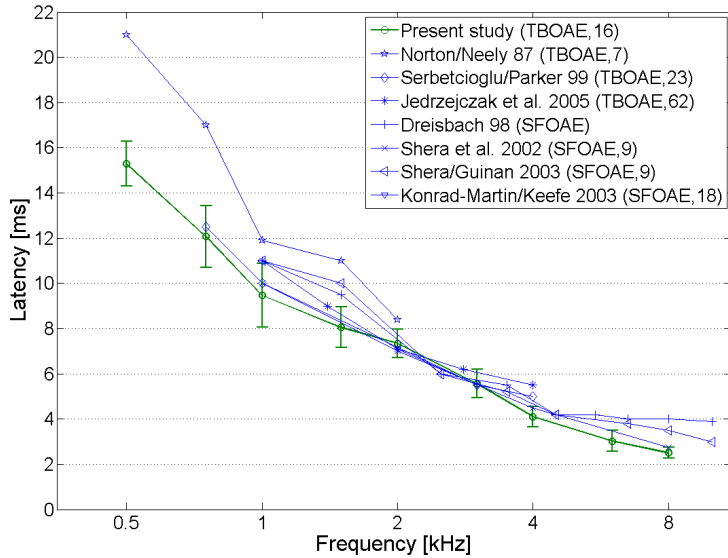


Figure 3.13: Comparison of OAE latencies obtained in different studies. The name of the authors is followed by the year of publication, the method and the number of subjects used. NB: For Jedrzejczak et al. this is the number of ears tested. For the present study, the latency is obtained with the same SNR for all subjects, see table 3.2 for the considered values. Most of the studies mentioned here have a maximum number of averages or fail to mention the method chosen.

### 3.3.5 Comparison of TBOAE latency with results from previous studies

A comparison between the mean OAE delays found in the present study and in previous studies has been made and is presented in figure 3.13. The general trend of the OAE latency is very similar in all studies. Only the data from Norton and Neely (1987) deviate clearly. This might be due to the differences in the methods. They used 1024 averages and ensured that the background noise was 30 dB below the OAE level. One important difference is how they dealt with onset ambiguity, they zeroed the part of the recorded signal corresponding to the stimulus plus 2 ms. This might have led to a suppression of the OAE, which could explain why "the emissions were usually not well localized in time and their latencies could seldom be precisely defined" in their study (Norton and Neely, 1987, page 1870). The OAE delays obtained in the

present study are slightly lower than in other studies. If the maximum number of averages would be considered, as in most of the other mentioned studies in the figure, the present results would be inside the range of these previous studies.

As a conclusion, it can be said that the present study provide an estimate of the cochlear delay which is in agreement with previous studies. The method (SFOAE or TBOAE) used to derive the latencies does not seem to have an effect on the estimate. However, the paradigm followed in the present study appears to be easier and more straightforward to implement.

### 3.4 Discussion

In the present study, a new paradigm to resolve the OAE onset ambiguity in the time domain was applied. The choice of the time domain to analyse the responses kept the processing quite simple and this new method led to results that agree with previous studies. Individual cochlear delays were estimated by using tone-bursts OAEs from 16 subjects. As expected, TBOAE latencies reflect the tonotopic mapping of the cochlea, with low frequencies TB having longer latencies than high frequencies. The values obtained range from around 16 ms at 0.5 kHz down to 2.5 ms at 8 kHz. These values are different for each of the 16 subjects employed. The inherent differences between subjects can be as high as 4 ms (at 1 kHz) whereas the latency estimates within one subject is highly repeatable (1.2 ms difference between recording sessions for the worst case). Different reasons for the variability of the latency estimates across subjects can be suggested: the first one is the reading error made while detecting the OAE, this difficulty is different for each subject, affecting the inter-subject variability. Another reason is related to the spectral splatter of the stimulus. The tone bursts used in this experiment not only excite the region of the tone-burst frequency but also neighbouring frequencies, as figure 3.1 showed. The secondary lobes of the stimulus spectrum are situated in the vicinity of the TB frequency, corresponding to regions of the cochlea around this frequency. For each subject, the BM excitation pattern is different, the secondary lobes lead to earlier excitation in some cases, meaning shorter latencies, hence a higher inter-subject variability. Wit and Ritsma (1980) also drew this conclusion.

The effect of the noise on the inter-subject variability was investigated. Fixing the SNR for all the subjects at each frequency led to slightly more reliable estimates of the TBOAE latency, with a slightly lower value than obtained with a maximum number of averages.

An observation that remains a matter of discussion is the non-monotonic decrease of the TBOAE latency estimates as a function of frequency, as observed in Fig. 3.9. The present study has only used information in the time domain, there could be a gain if the frequency domain was investigated too. A time-frequency representation allows to see what is happening in the time domain and frequency domain at the same time. This would give more information about the OAE burst onset and cues about the presence of spontaneous OAE (SOAE). An example of the time-frequency analysis for one subject is shown in figure 3.14, this representation is obtained via a spectrogram calculated using a window of 664 samples with an overlap of 500 samples. The FFT was calculated from 10 Hz to 8 kHz with a sampling frequency of 48 kHz.

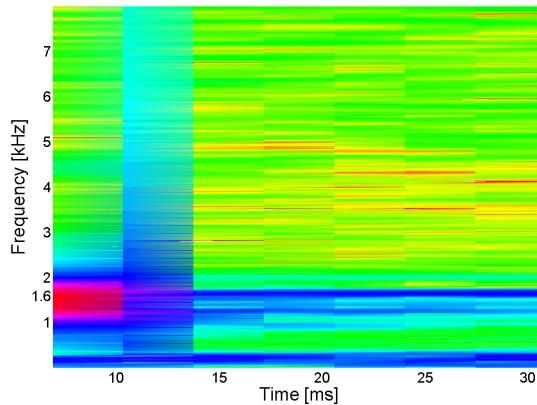


Figure 3.14: Spectrogram of a recorded responses evoked by a 1.5 kHz tone burst. The red straight trace at around 1.6 kHz confirms that there is a continuous ringing throughout the recording. This is believed to be a spontaneous OAE.

The red patch on the left of the spectrogram represents the stimulus. The most striking phenomenon that can be seen in this figure is the almost straight blue mark at around 1.6 kHz. This implies a continuous ringing at this frequency along

the recording. This suggests that the subject has spontaneous OAEs. The latencies obtained for this specific subject are shown in figure 3.15.

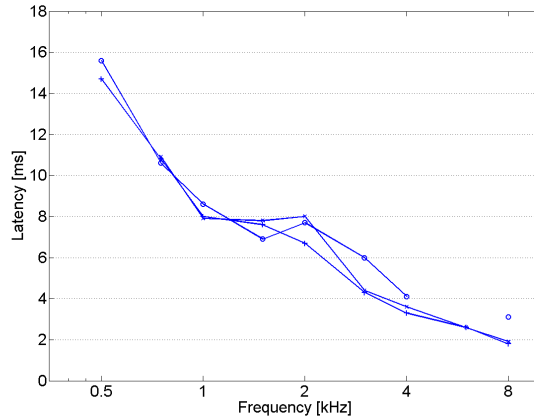


Figure 3.15: Latency estimate for subject 13 for three different sessions. Note the plateau between 1 and 2 kHz, the latency does not seem to change much, this has to be related with the time-frequency plot shown in Fig. 3.14

The latency curves have an unusual behaviour around 1-2 kHz, with a local maximum at 2 kHz. This effect as well as the difficulty to discern the correct OAE burst can now be explained by the presence of SOAEs around 1.6 kHz. This time-frequency analysis was performed for subjects whose latency as a function of frequency was a non-monotonically decreasing function. The presence of SOAE was found in most of these cases, usually occurring between 1 and 2 kHz. Three out of the four female subjects demonstrated SOAEs, confirming the prevalence observed in previous studies (Probst et al., 1986). However, for some subjects, no SOAEs could be found. In these cases, the "knee" could reveal the actual state of the cochlea. Perhaps, for these subjects, the BM filters around 2 kHz are larger than for other subjects and lower frequencies such as 1.5 kHz are processed by this filter instead, this irregularity could then explain the plateau observed. This plateau would also appear in the estimate of evoked potentials latency as a function of frequency.



### 3.4.1 Conclusion

The comparison with previous studies has shown that the current paradigm used to obtain TBOAE is reliable. It presents a good alternative to other methods to estimating the cochlear delay. This method has the advantage that only time-domain information is used and the post processing is kept simple.

Due to the great inter-subject variability observed at low frequencies, it seems more justified to consider individual subject measures instead of averaging across subjects. Indeed, the audibility and cochlear status of each individual vary quite a lot. Further investigation would be needed to explain the "knee" observed in some subjects in the absence of SOAE. Also, the precise effect of the middle ear on OAE could be searched in a future phase. By compensating for a possible filtering effect of the middle ear, the OAE latency could be more accurately estimated. In order to assess the robustness of the results achieved in this study, it would be valuable to estimate cochlear delay with an alternative non-invasive method. Auditory evoked potentials could provide a good candidate for such experiments.



# Individual estimates of brainstem response delays

---

## Abstract

Chapter 3 presented a method to estimate the latency of OAEs evoked by tone bursts in individuals. The current chapter uses another physiological method to estimate the latency of the wave V in ABR recordings. The 11 subjects of this experiment are a subgroup of the group that took part in the OAE experiment. This study also assesses the degree of inter-subject variability. The robustness of the measures is demonstrated via repeat recordings and the associated intra-subject variability. This ABR experiment shows how an estimate of the cochlear delay can be obtained from the wave-V latency. For the individual subjects, the delay at each tone-burst frequency is reproducible. The wave V delays are in good agreement with previous research. When differences are observed with other published studies, they can be accounted for by the stimulus and level difference between studies. The conclusion of the measurements presented in this chapter is that cochlear delays estimated from reliable ABR measures can be compared with the previously estimated OAE delays.

## 4.1 Introduction

The experiment presented in Chap. 3 was reliable with a low intra-subject variability, it showed a higher inter-subject variability at low frequencies and highlighted a non-monotonic decrease of the OAE latency as a function of frequency observed for some subjects. As seen in chapter 2, there exist other methods to estimate

auditory delays, among which is the well established auditory brainstem responses recording technique. Although this technique has mainly been used to estimate auditory threshold (Stapells and Picton, 1981), it is possible to indirectly obtain BM latency information as well. In this chapter the experiment to obtain ABR latency estimates is described. The direct comparison with OAE delays will be presented in chapter 5.

In order to have comparable results with the OAE experiment, it is necessary to also use frequency-specific stimuli. As explained in chapter 2, there exist different techniques to elicit frequency-specific ABR, such as the masking or tone bursts (TB) techniques. Tone bursts will be used here due to the experimental simplicity. Many studies have used TB to estimate wave-V latency (Kodera et al., 1977a; Suzuki et al., 1977; Suzuki and Horiuchi, 1977; Burkard and Hecox, 1983; Maurizi et al., 1984; Davis et al., 1985; Gorga et al., 1988; Conijn et al., 1990; Fausti et al., 1993; Beattie et al., 1994; Beattie and Torre, 1997; Oates and Stapells, 1997; Murray et al., 1998). However, most of these studies investigated a limited frequency range. The most extensive study was done by Gorga et al. (1988) who investigated TBABR latencies with ten TB centre frequencies between 0.25 and 8 kHz.

The present experiment aims at comparing wave V latencies to the OAE latencies estimated in chapter 3. That is why the stimuli chosen for this experiment will be similar to the OAE study, i. e. tone bursts ranging from 0.5 to 8 kHz. Recording tone-burst ABR at 250 Hz has proved to be too noisy to give usable results.

In the following, the stimuli used and the test persons involved in the experiment are detailed. Then, similar to the OAE experiment in Chap. 3, the intra- and inter-subject variabilities are investigated. It is shown that wave V latencies have good test-retest reliability for each subject, and suggestions are made to explain why the inter-subject variability is frequency dependent. This chapter also presents the problem of estimating cochlear delays from wave-V latency estimates. Finally, a comparison with data previously reported in the literature will be presented.

## 4.2 Methods

### 4.2.1 Subjects

The eleven subjects participating in this experiment were a subgroup of the 16 normal-hearing adults used in the OAE experiment. There were 2 females and 9 males, all having pure-tone thresholds better than 15 dB HL in the range 0.25-8 kHz. The possible effect of the higher number of male subjects on the results will be discussed later. Each subject was paid for their participation. They were asked to lay down on a clinical couch in a sound insulated booth, and were asked to relax as much as possible, where most fell asleep. A state of sleep in the subject showed a reduced background noise and does not influence ABR amplitude or latency (see chapter 2). The electrophysiological recordings were collected during three sessions, lasting about two hours each. Subjects were offered to take a break at the middle of the session. Post processing was done to obtain the ABR averages.

### 4.2.2 Stimuli

The stimuli used in this experiment had the same characteristics as the tone bursts used in the OAE experiment. Their properties are summarized in table 3.1 page 26. The length of these stimuli is the result of a compromise between the need for a rise time constant in milliseconds and the need for a rise time constant in number of cycles. The stimuli are of ratio type 1:0:1 (rise:plateau:fall) as suggested by Stapells (1994). As table 3.1 shows, the rise times are longer for lower frequencies, 5 ms at 0.5 kHz and 0.625 ms at 8 kHz. It is known that the rise time is responsible for the neural synchrony leading to the brainstem responses (Suzuki and Horiuchi, 1981). To obtain a detectable response it is necessary to have a sharp stimulus onset (i. e. a short rise time) but not too short so as to keep a good frequency specificity. A very short or transient signal has a broad frequency spectrum. And on the other hand, an increase of this rise time makes the stimulus bandwidth narrower, high frequency regions of the BM are therefore not excited, which leads to an increase of the wave-V latency (Stapells and Picton, 1981; Beattie and Torre, 1997). This is an inherent drawback to accept in order to keep the

stimuli frequency-specific. The influence of this rise time difference between high and low frequencies on wave-V latency is discussed later in appendix E.

### 4.2.3 Procedure

A preliminary experiment showed that alternating the polarity of the presented stimuli reduces the effect of a stimulus artifact. Despite the shielding of the transducer, some electrical signal due to the stimulus could be recorded in the response and the alternating polarity procedure yielded the best results. It has previously been shown that this also leads to higher amplitude of the wave V, making its detection easier (Fuxe and Stapells, 1993; Schönweiler et al., 2005). The preliminary study also showed that the software used to record the responses would crash unexpectedly if too high repetition rates were chosen. A good compromise was found with 24.5 repetitions per second. Studies have shown that the amplitude of the responses remains unaffected by repetition rates up to 35/s (Stapells and Picton, 1981) and the latency is stable for a large range of repetition rates (Burkard and Secor, 2002).

The electrodes placement of this experiment follows the 10-20 electrode system suggested by Jasper (1958). An example is shown in figure 4.1. The electrode placed at the forehead (Fz) is used as the ground electrode, the one at the vertex (Cz) is used as reference and the two electrodes on the mastoid, M1 and M2, record the signal at the ipsilateral and contralateral mastoid respectively. The impedance of each electrode was measured prior to the measurement and was below the commonly used limit of 5 k $\Omega$ .

The stimuli were calibrated at 66 dB peSPL, the calibration procedure used for brief tones is described in appendix B. The order of presentation of the frequencies was randomized for each subject. Each recording session started with a control measurement in order for the subject to relax while getting used to the setup and the stimulus level. The stimulus used was the flat-spectrum chirp from Dau et al. (2000) presented 3000 times. These recordings were not analysed further. The number of averages used for the tone bursts varied with frequency. A preliminary experiment showed that the wave V could be detected using fewer averages for high frequencies, due to the stronger signal strength at these frequencies. The lower frequencies (0.5, 0.75, 1 kHz) were repeated 8000 times, the middle frequencies (1.5, 2, 3 kHz) 4000

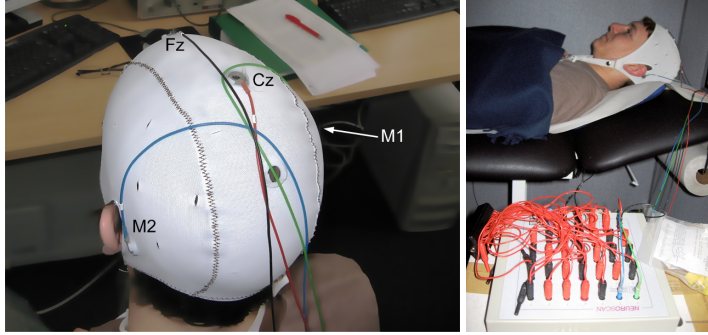


Figure 4.1: (Left) The subjects wore a cap on which the electrodes were positioned: forehead (Fz), vertex (Cz), ipsilateral (M1) and contralateral (M2) mastoid. (Right) Subject lying on a couch with earprobe in the ear and electrodes connected to the amplifier.

times and the higher frequencies (4, 6, 8 kHz) 3000 times for each run. One session consisted of two runs and lasted approximately two hours.

The stimuli were produced by a computer using Matlab, then sent to a D/A converter (RME ADI8-Pro). The analog signal was transmitted to a programmable attenuator (TDT PA5) and a headphone driver (TDT HB7) and finally to the insert earphone ER-2 (see figure 4.2). The responses were recorded using the Synamp amplifier and the accompanying software from Neuroscan Inc. The data were saved on disk for off-line analysis, they were then epoched and averaged using an iterative weighted-averaging algorithm (Riedel et al., 2001). The responses were filtered then, between 0.1 and 1.5 kHz.

#### 4.2.4 TBABR latency and interpeak delays

The responses to the nine frequencies were plotted on the same graph to help identifying the different peaks. To recall from chapter 2, the wave-V latency can be defined as the sum of three delays: the cochlear delay,  $\tau_{BM}$ , the synaptic delay,  $\tau_{synaptic}$  and the neural delay,  $\tau_{neural}$ . The separation of these delays is schematized in figure 4.3.

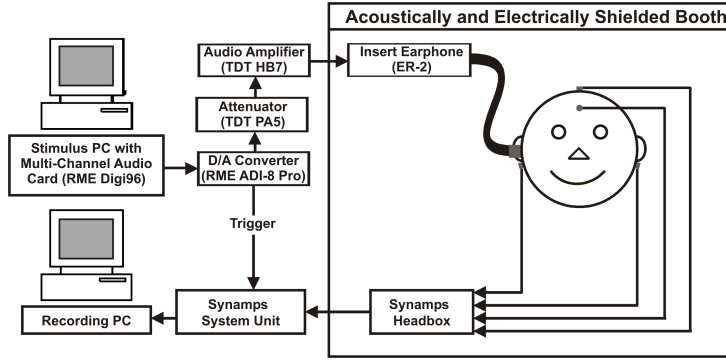


Figure 4.2: Equipment used for the recording of ABR. The stimuli were produced by a computer using Matlab, then sent to a D/A converter (RME ADI8-Pro). The analog signal was transmitted to a programmable attenuator (TDT PA5) and a headphone driver (TDT HB7) and finally to the insert earphone ER-2. The responses were recorded using the Synamp amplifier and the accompanying software from Neuroscan Inc.

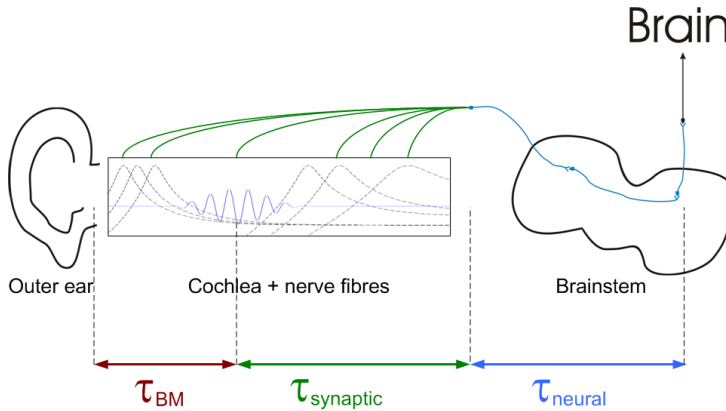


Figure 4.3: The latency of the wave V can be defined as the sum of three delays: the cochlear delay ( $\tau_{BM}$ ), the synaptic delay ( $\tau_{synaptic}$ ) and the neural delay ( $\tau_{neural}$ ).  $\tau_{BM}$  is the time between the stimulus onset and the activation of the nerve fibers,  $\tau_{synaptic}$  is the transport time of these nerve fibers and  $\tau_{neural}$  represents the time between the auditory nerve and the place responsible for wave V.



The exact measure of the cochlear delay is noted  $\tau_{BM}$  and its estimate from ABR measurement is noted  $\hat{\tau}_{BM}^{(ABR)}$ . Following the definition of the wave-V latency:

$$\begin{aligned}\tau_{wave\ V} &= \hat{\tau}_{BM}^{(ABR)} + \tau_{synaptic} + \tau_{neural} \\ \hat{\tau}_{BM}^{(ABR)} &= \tau_{wave\ V} - \tau_{neural} - \tau_{synaptic}\end{aligned}$$

The cochlear delay can thus be obtained by subtracting the synaptic and neural delays to the wave-V latency. The synaptic delay is usually approximated around 1 ms (Kiang, 1975; Kim and Molnar, 1979; Neely et al., 1988; Burkard and Secor, 2002), independently of frequency and level (Don et al., 1998). The neural delay can be estimated from the interpeak delays, of the ABR. Wave I is believed to arise at the extremity of the auditory nerve, at the junction between the cochleo-vestibular nerve and the brainstem (see section 2.2.1). The neural delay can therefore be estimated by the latency difference of wave V and wave I, i. e.  $\tau_{neural} = \Delta_{I-V}$ . However, unlike wave V, the detection of wave I is rather difficult. It was therefore decided to detect wave III instead and use the assumption  $\Delta_{I-V} = 2 \Delta_{III-V}$  (Don and Eggermont, 1978; Eggermont and Don, 1980; Don and Kwong, 2002). The cochlear delay estimate  $\hat{\tau}_{BM}^{(ABR)}$  can then be calculated for each subject, using individual estimate of wave-V latency and interpeak delay:

$$\hat{\tau}_{BM}^{(ABR)} = \tau_{wave\ V} - 2 \Delta_{III-V} - 1 \text{ ms} \quad (4.1)$$

## 4.3 Results

### 4.3.1 wave-V latency

An example of ABR recordings for one subject is presented in figure 4.4. Each pair of curves represents one tone-burst centre frequency (see vertical axis), the solid and dashed line are two runs of the same session to show repeatability and help identifying the prominent peak. The peaks corresponding to the wave V appear clearly at each frequency and are marked by 'V'. At high frequencies (Fig. 4.4 (left)), it is possible to detect wave III, marked by '↓'. The delay between wave III and wave V will be used to estimate  $\tau_{BM}$ .

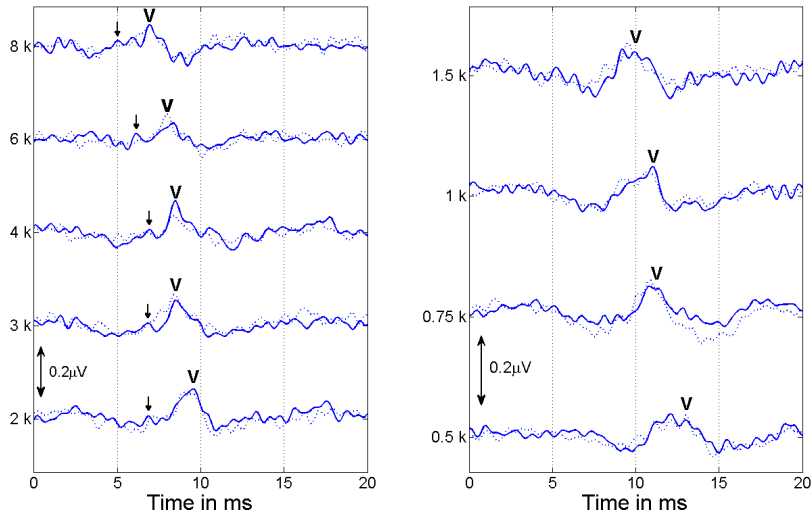


Figure 4.4: ABR recordings for one subject at high and low frequencies. Each pair of curves represents one frequency and each curve is a run (solid blue line is run 3 and the dashed line is run 4). The wave V appears clearly at each frequency and is marked by 'V'. Wave III can be detected for high frequencies and is marked by '↓'. Frequencies are in Hz.

The general trend is an increase of the wave-V latency with decreasing frequency, from around 7 ms at 8 kHz to around 12 ms at 0.5 kHz. There is, however, an ambiguity regarding the attribution of the correct peak to wave V. This is particularly evident in this example for 0.5 and 6 kHz. An important issue was the detection of the wave V at low frequencies, it turned out to be impossible for some subjects to detect the wave V despite the high number of averages (8000). This problem at low frequencies has been pointed out in previous studies (Stapells, 1994; Stürzebecher et al., 2006) and could be due to the lower speed of the travelling wave in the low-frequency region of the cochlea compared to the basal part. With a lower velocity, adjacent nerves fire with a certain delay, leading to asynchronous neural activity, thus the general level of the ABR is lower. An increase in the level of the stimulus would definitely give a more detectable wave V at 0.5 kHz, due to a greater discharge of

the neurons (Gorga et al., 1988). But the stimuli level of the TBABR measurements (66 dB peSPL) was chosen according to the OAE experiment for better comparability. A level higher than 66 dB peSPL in the OAE experiment would have resulted in a reduced active mechanism of the cochlea and therefore a reduced frequency specificity due to broader filter activation (Ruggero et al., 1997).

Stapells (1994) gives a number of suggestions to improve the wave V detection at low frequencies. He suggested, for example, to lower the high-pass filter from 100 Hz to 30 Hz in an attempt to collect more signal. This was tried and did not improve the readability of the low frequencies ABR significantly, more noise was actually added in the filtered EEG. The ambient noise level increases as frequency decreases (Stapells, 1994), i. e. low frequencies ABR are more susceptible to ambient noise than high frequencies ABR. But the measurement being done in a certified insulated booth, it is not believed that the environment is responsible for the higher noise level. It is possible that the ABR recordings are polluted by physiological noise from the subject, not all of them were asleep during the recording. Another possible source of noise during the recording is the signal amplifier, positioned between the electrodes and the recording PC.

#### **4.3.2 Intra- and inter-subject variability**

The estimation of the wave-V latency for all eleven subjects is shown in figure 4.5. As expected the wave-V latencies decrease with increasing frequency. For some subjects this trend is non-monotonic and no clear explanation was found for this behaviour. However, to clarify this phenomenon, a comparison for individual subjects between the latency of the TBABR and the TBOAE is presented in chapter 5. Another general comment is that more data were collected at high frequencies, making the wave-V latency more reliable at these frequencies, this is due to the easier detection of wave V in this range.

The intra-subject variability is presented in the figure by  $\pm 1$  standard deviation (error bars) and can be found in appendix D. A clear observation is that it is greater at low frequencies. The highest intra-subject variability is reached for subject 3 at 0.5 kHz with a standard deviation of 1.13 ms. This difference of latency for the same subject at the same frequency can be partly explained by the difficulty of assessing the

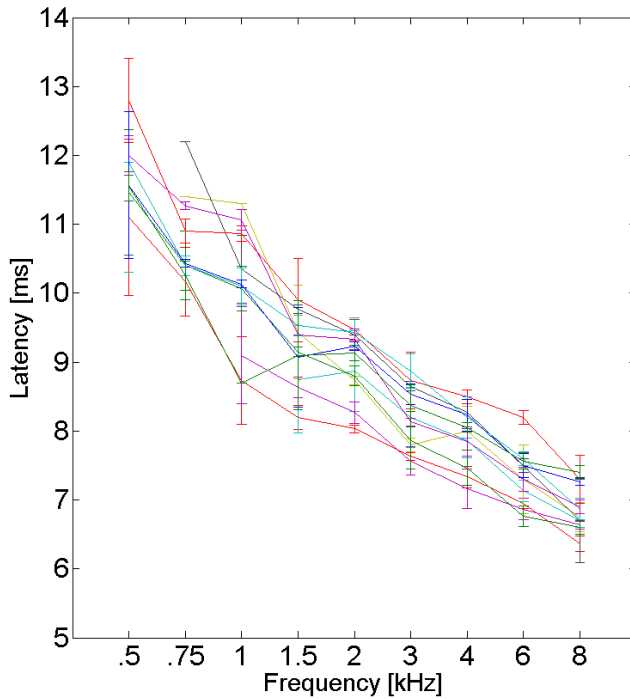


Figure 4.5: wave-V latency estimation for all eleven subjects at nine frequencies. The intra-subject variability is indicated by the error bars ( $\pm 1$  std), it is greater at low frequencies. It also shows the inter-subject variability which appears greater at low frequencies.

correct peak to the wave V. This maximum standard deviation of 1.13 ms reveals that the precision reached in the latency estimate is rather good if one takes into account the processing involved in wave V generation.

Figure 4.5 also demonstrates the inter-subject variability, which appears greater at low frequencies. Previous studies also reported larger inter-subject variability at low frequencies (Gorga et al., 1988; Murray et al., 1998). As explained in chapter 2, the latency of the ABR is affected by several parameters such as age, head size, body core temperature, gender and physiological noise, these parameters might account for the inter-subject variability.

There are different reasons to explain the lower intra- and inter-subject variability at high frequencies compared to low frequencies. The first explanation is related with the difference in stimulus rise time. As mentioned earlier (section 4.2.2), the stimulus rise time is related to the neuronal discharge synchrony. A higher discharge synchrony produces ABR with higher amplitudes. At equal rise times, high frequency TB have narrower bandwidth than low frequencies TB, this leads to an increases of the ABR amplitude with frequency (Stapells and Picton, 1981; Schönweiler et al., 2005). This trend is even more accentuated if the low frequencies stimuli have longer rise times than the high frequencies stimuli, for these latter the discharge synchrony is higher, leading to greater ABR amplitude. But this choice of different rise time is necessary to have frequency-specific stimuli. The consequence of the rise time is therefore that waves V are easier to detect in the background noise for the higher frequencies and this leads to lower intra- and inter-subject variabilities. Another reason could be the phase locking occurring at low frequencies. The nerve fibers fire at a rate depending on the tone-burst frequency and this happens roughly at the same phase (Moore, 2003). A change in the stimulus phase, e. g. by alternating polarity, could introduce a jitter in the responses.

### 4.3.3 Interpeak intervals

As mentioned in section 4.2, the intervals between the different ABR peaks are of interest to find the estimate of the cochlear delay  $\hat{\tau}_{BM}^{(ABR)}$ . Equation 4.1 requires the time interval between wave III and wave V. Since the waves are more easily detected at high frequencies and the interpeak delay is independent of frequency (Don and Eggermont, 1978; Eggermont and Don, 1980), it was decided to estimate  $\Delta_{III-V}$  at frequencies above 2 kHz. The interpeak intervals are shown in table 4.1. The interpeak delay,  $\Delta_{III-V}$ , is roughly constant across frequencies for a given subject and in good agreement with values found previously (Don and Eggermont, 1978, Fig. 6).

As the present study is aiming at *individual* estimates of the cochlear delay, it is necessary to have an *individual* estimate of the interpeak delay. The differences in interpeak delay between the eleven subjects can be as large as 0.5 ms. There would be a loss of information by taking an average across subjects, as did most of the previous studies.

Frequency [kHz]	2	3	4	6	8	mean [ms]
Subject 3	1.6	1.7	1.7	1.8	1.7	<b>1.7</b>
Subject 4	1.6	1.8	1.6	1.5	1.6	<b>1.58</b>
Subject 5	1.9	2.1	2.2	2	2.2	<b>2.08</b>
Subject 6	1.9	2	2	2	2	<b>1.98</b>
Subject 7	1.8	1.7	1.9	1.8	1.7	<b>1.78</b>
Subject 9	1.6	1.8	1.7	1.6	1.5	<b>1.7</b>
Subject 10	1.9	1.7	1.6	2	1.9	<b>1.82</b>
Subject 11	1.6	1.7	1.4		1.6	<b>1.57</b>
Subject 12	1.6	1.5	1.7	1.5	1.7	<b>1.6</b>
Subject 15	1.7	1.6		1.7	1.7	<b>1.67</b>
Subject 16	1.7	1.6	2	1.6	1.8	<b>1.78</b>

Table 4.1: Interpeak delay estimation,  $\Delta_{III-V}$  for the eleven subjects participating in the ABR experiment. As reported in previous research, the interpeak delay is rather constant across frequency (Don and Eggermont, 1978) but some variations across subjects can be observed. The subject identification number is the same as for the OAE experiment. The blanks indicate that no values could be found.

## 4.4 Discussion

A comparison between the latency of wave V estimated in the present study and an extensive list of published data is presented in figure 4.6. Care must be taken when comparing different studies. For most of them the measure of latency is not the main aim. Some investigated the effect of SNR (Burkard and Hecox, 1983; Beattie et al., 1994) or stimulus rise-time (Beattie and Torre, 1997) on ABR. Others examined the use of ABRs as an audiometric tool (Kodera et al., 1977a; Stapells and Picton, 1981; Conijn et al., 1990; Murray et al., 1998). These studies sometimes fail to mention the complete procedure used, such as stimulus length or level in dB SPL. It is however possible to extract data from some of them and compare their results with the present study in a meta-analysis. All the results presented here used tone-bursts to measure ABR for normal hearing test persons. The level used is given in the plot after each study. Most of the studies reported the level in dB nHL or dB SL, it was sometimes possible to find a correspondence of the level in dB peSPL in the paper itself or in accompanying papers. Based on these values and on the fact that  $0 \text{ dB HL} \approx 10 \text{ dB SPL}$  for pure tones at the frequencies considered (ANSI, 1973; ISO, 1994), a factor was found to transform dB nHL into dB SPL for tone-bursts:

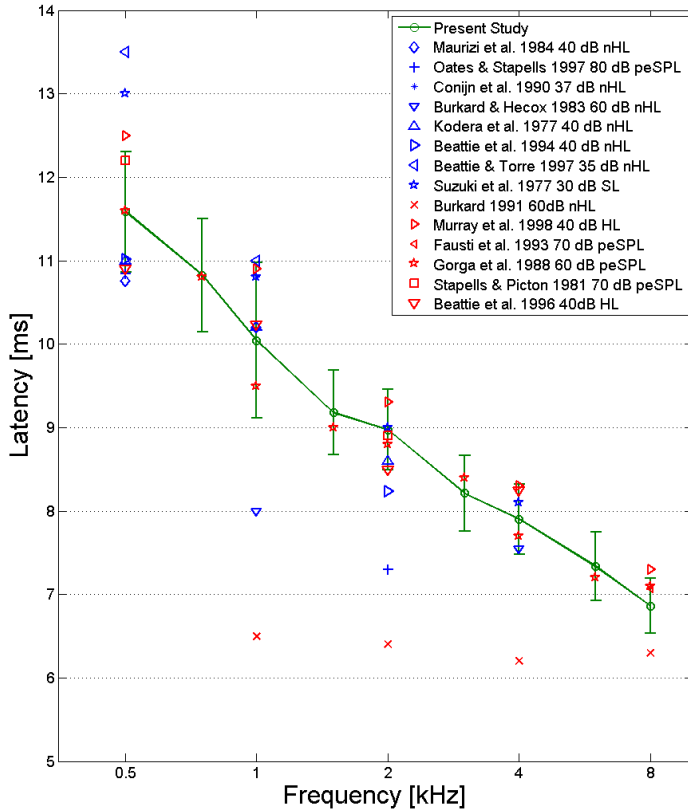


Figure 4.6: Comparison of wave-V latency with previously reported data. The green error bars indicate  $\pm 1$  std of the wave-V latency estimated in the present study, this is an indication for the inter-subject variability. The present data were collected at 66 dB peSPL or 38.8 dB HL.

0 dB HL  $\approx$  25 dB SPL. The same factor was used by Gorga et al. (1988). With this factor and the hearing level of each subjects, the level used for the present experiment (66 dB peSPL) is equivalent to 38.8 dB HL.

The green line links the data from the current study, the standard deviation is indicated with the error bars. The standard deviation reflects the inherent inter-subject variability. It was decided not to plot the variability of other studies for clarity. A first glance at figure 4.6 leads to the conclusion that there is a good agreement with previous data. The discrepancies between the studies can be explained by three main

reasons: 1) the rise time of the stimuli, 2) the stimulus level and 3) the gender of the subjects.

#### 4.4.1 Rise time difference

Stapells and Picton (1981) showed that, for each frequencies, a decrease of the rise time leads to a lower wave-V latency. The rise times used in this study are presented in table 4.2. The different rise times across frequencies is explained by the need for frequency specificity.

Frequency [kHz]	0.5	0.75	1	1.5	2	3	4	6	8
Rise Time [ms]	5	3.5	2.5	2.5	2.5	1.7	1.25	0.85	0.625

Table 4.2: Rise time of the different tone-bursts. As explained in a previous section, the need for frequency specificity leads to these different rise times.

Most of the studies presented here used different rise times (Kodera et al., 1977a; Maurizi et al., 1984; Conijn et al., 1990; Burkard, 1991; Beattie et al., 1994, 1996; Oates and Stapells, 1997; Murray et al., 1998). This difference can be as large as 4 ms (Beattie et al., 1994) which, according to Stapells and Picton (1981) Fig. 8, affects the wave-V latency by approximately 1.6 ms. Based on the values given by Stapells and Picton (1981), it is possible to compensate for the difference in rise time and to explain some of the outliers in figure 4.6. For example, Maurizi et al. (1984) used a rise time of 1.5 ms at 0.5 kHz, compensating for the rise time difference with the present study shifts this data point by 1.4 ms upwards, placing it within the range of the present study. The same applies for Burkard (1991) and Beattie et al. (1994, 1996).

#### 4.4.2 Level difference

The stimulus level is also an important parameter. As seen in Chap. 2, an increase in level leads to a decrease of the wave-V latency (Gorga et al., 1988). Some studies mentioned here used stimulus levels which differ by up to 23 dB (Burkard, 1991, at 1 kHz). Based on figure 5 from Gorga et al. (1988), it is possible to compensate for such a level difference. For instance, the study from Beattie and Torre (1997) at 0.5



and 1 kHz uses levels of 58.9 and 51.5 dB peSPL respectively. Compensation factors of  $-0.8$  ms at 0.5 kHz and  $-1.5$  ms at 1 kHz were found and if their large variability (1.9 and 1.7 ms respectively) is taken into consideration, their values agree with the present study. The difference in levels could also account for other outliers observed in figure 4.6, such as Burkard and Hecox (1983); Burkard (1991) and Oates and Stapells (1997). For some studies, both a difference in rise time and level explain their outside position, e. g. Burkard (1991).

#### **4.4.3 Subject gender difference**

Beattie et al. (1994) recorded ABR on 15 female subjects at 0.5 and 2 kHz. As seen in Chap. 2, female subjects have shorter wave V latencies than males. This different gender ratio with the present study (9 males, 2 females) could explain, to some extent, the lower latencies found by Beattie et al. (1994).

From the comparison with results previously published, it can be concluded that the results achieved in this study are reliable and that differences in stimulus rise time and level can account for a great deal to the discrepancies observed.

#### **4.4.4 Summary**

This chapter has shown that the ABR experiment carried out on 11 subjects led to very reliable results. The estimate of the wave-V latency remains stable across measurements for each subject, with a variability lower than 1.13 ms. The variability across subjects reflects the inherent differences from person to person. One of the reason could be the sensitivity of the recording for background noise. The results obtained during this experiment are very similar to previously published data, this is true for both the mean of the wave-V latency and the inter-subject variability. These reliable data can now be compared to the, also reliable, OAE data recorded earlier.



## Comparison of OAE and ABR estimates of cochlear delay

---

*This work was presented at the International Symposium on Auditory and Audiological Research (ISAAR), 29-31 August 2007, Helsingør, Denmark and published in the proceedings "Auditory Signal Processing in Hearing Impaired Listeners" edited by T. Dau, J. Buchholz, J. Harte and T. Christiansen.*

### Abstract

Estimates of OAE delays were obtained in chapter 3 and an estimate of the basilar membrane delay,  $\hat{\tau}_{BM}^{(ABR)}$ , can be obtained from the wave-V latency, as described in chapter 4. The present chapter compares the BM latency estimate,  $\hat{\tau}_{BM}^{(ABR)}$ , with the OAE latency,  $\tau_{OAE}$ . The comparison is based on the eleven subjects that participated in both experiments. The observed latencies as a function of frequency are qualitatively similar across subjects. For the individual subjects, the delay at each tone-burst frequency is reproducible. A difference in inter-subject variability between TBOAE and TBABR is apparent at low frequencies. Attempts are made to understand this in the individuals tested. Three current theories about the generation and propagation of OAEs are presented, whereby each theory implies a different relation between  $\hat{\tau}_{BM}^{(ABR)}$  and  $\tau_{OAE}$ . These relations are confronted with the experimental results of this study. The acoustic wave theory, the CRF theory and the signal-front theory do not appear to hold over the entire frequency range. Based on this, suggestions are made about the possible generation sites of OAEs and ABRs and theoretical implications on the transmission of the travelling wave are discussed.

## 5.1 Introduction

Although the origins of ABRs are fairly well understood, the generation mechanism of OAEs is still in question (see Chap. 2). The exact relation between the OAE delay,  $\tau_{OAE}$ , and the BM delay estimated from ABR,  $\hat{\tau}_{BM}^{(ABR)}$  in Eq. 4.1, remains a source of discussion. Since the discovery of OAEs, different theories have emerged to explain their origin and their mode of propagation. This section introduces theories implying different relations between  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$ . Each of these two delays can be expressed in terms of a power law and can be interpreted as being proportional to the frequency raised to some negative power less than unity, i.e.  $\tau = bf^{-a}$  where  $a < 1$ . Plotting  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  on a loglog axis means that any such power law will be observed as a straight line with slope  $-a$ . The relation between  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  is hence reflected in the relative position of their curve.

### 5.1.1 Acoustic wave hypothesis

The acoustic wave hypothesis (also called fast wave theory) has been explained in Ruggero (2004) and was first suggested by Wilson (1980b). It suggests that OAEs are generated at the place where the BM transfer function reaches its peak and that, in this region, the OHC undergo a volumetric change, a "synchronous swelling and shrinking" (Wilson, 1980b). A compressional wave is then created that propagates through the cochlear fluid rather than through the cochlear partition (membrane) to cause the stapes to vibrate. The propagation of the compressional wave is very fast in the fluid with respect to the forward travelling time. This leads to a negligible backward travel time of the OAE signal. In this case, the latency of the OAE is predicted as being  $\tau_{OAE} = \tau_{BM}$ , as indicated in Fig. 5.1.

### 5.1.2 Coherent Reflection Filtering theory

The CRF theory was suggested by Zweig and Shera (1995) and is widely accepted as a description of the cochlear mechanics. It states that OAEs originate from coherent reflection due to random impedance perturbations mostly from the sites of maximum excitation of the BM (see chapter 2). The travelling wave propagates backward along

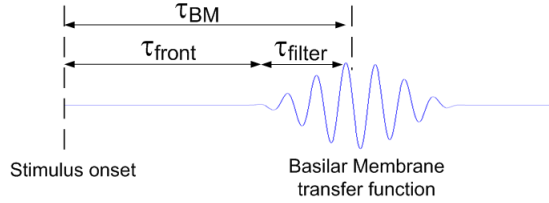


Figure 5.1: The delays measured by different non-invasive methods (e. g. OAE, ABR, ECochG) are the sum of delays corresponding to various processing along the auditory pathway. The exact relation between the OAE delay,  $\tau_{OAE}$ , and the BM delay,  $\tau_{BM}$ , remains the source of numerous discussions.  $\tau_{BM}$  is suggested to be the sum of the signal-front delay of the inward travelling wave,  $\tau_{front}$  or  $\tau_{transport}$  and of the cochlear filter delay,  $\tau_{filter}$ .

the BM and causes the oval window to vibrate. This results in the signal recorded in the ear canal. This round-trip theory suggests that the travel time for the OAEs is twice the travel time to the CF location on the BM, i. e.  $\tau_{OAE} = 2\tau_{BM}$ . This assumption of an equal travel time in both direction has been used in many studies since its publication (Serbetçioğlu and Parker, 1999; Schoonhoven et al., 2001; Goodman et al., 2004). But already 15 years before Shera and Zweig, Rутten (1980) suggested the same hypothesis and the results from Neely et al. (1988) were consistent with that idea. There are also several studies that challenge the round-trip assumption, such as Nobili et al. (2003), Ren (2004), Siegel et al. (2005) and He et al. (2007). The experimental data of the present study will help clarifying the different views on the CRF theory.

### 5.1.3 Signal-front hypothesis

The signal-front hypothesis was first suggested by Ruggero (2004) (hypothesis IA). This hypothesis agrees with the backward travelling wave generated at the site of maximum BM excitation, as explained by the CRF theory. However, the signal-front hypothesis differs from the two previous hypotheses by the calculation of this backward travel time. It divides the OAE delay into two components: the signal front delay,  $\tau_{front}$ , which corresponds to the time between the stapes vibration and the BM vibrations, and the BM filter delay,  $\tau_{filter}$ , see figure 5.1.  $\tau_{front}$  is also the  $\tau_{transport}$  presented in Chap. 2. According to the signal-front theory, the OAE delay is a round-

trip between the stapes and the start of the BM oscillation, to which the BM filter delay is added. The relation can be written as:

$$\begin{aligned}
 \tau_{OAE} &= 2\tau_{front} + \tau_{filter} \\
 &= 2\tau_{front} + (\tau_{BM} - \tau_{front}) \\
 &= \tau_{front} + \tau_{BM}
 \end{aligned}$$

These different theoretical relations between  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  are typically based on animal data; only the work from Wilson (1980b) includes human data. There is therefore a need to confront each of the three theories with experimental data in humans. Such data have been collected in this study.

## 5.2 Results

This section first compares the trends of  $\hat{\tau}_{BM}^{(ABR)}$  and  $\tau_{OAE}$  and their respective intra- and inter-subject variabilities. It then confronts each of the aforementioned theories with the current data.

As expected, OAEs and ABRs show exponentially decreasing delays with increasing frequencies. The rate of this decrease and how OAEs and ABRs compare depend on their generation mechanisms. Figure 5.2 shows the estimates  $\hat{\tau}_{BM}^{(ABR)}$  and  $\tau_{OAE}$  as a function of frequency for two illustrative subjects. The curves for all subjects are reported in appendix F.  $\hat{\tau}_{BM}^{(ABR)}$  was calculated from the wave-V latency following equation 4.1. The OAE curves are above the ABR curves, indicating longer latencies for OAEs than for ABRs. The OAE curves also have a steeper slope than the ABR curves. Besides, the ABR data show a smaller range of latencies. This observation is not so startling since OAEs and ABRs have clearly different travelling path, the exact relation between these travel times has yet to be defined and will be discussed further below. An observation made in the OAE study was the appearance of a non-monotonically decreasing  $\tau_{OAE}$ ; some plateaus were observed around 2 kHz for some of the listeners (see Fig. 3.15 on page 46). A time-frequency analysis showed that, for some subjects, spontaneous OAEs might be responsible for the plateaus. This conclusion is corroborated by the ABR measurement since  $\hat{\tau}_{BM}^{(ABR)}$  as a function of

frequency does not present this "knee". This is illustrated in figure 5.2 (left) where the activity seen around 1-2 kHz at a cochlear level ( $\tau_{OAE}$ ) is not present at a brainstem level. This observation confirms the presence of SOAEs, as suggested by the time-frequency analysis. On the other hand, this non-monotonic decrease of  $\tau_{OAE}$  was also observed in subjects where no SOAEs could be found, see Fig. 5.2 (right) for subjects 15. The variability of the ABR recordings at 1.5 kHz prevent from drawing a strong conclusion and the "knee" may or may not be present in the ABR data. The question regarding possible cochlear irregularities stays open and could be interesting to investigate in the future.

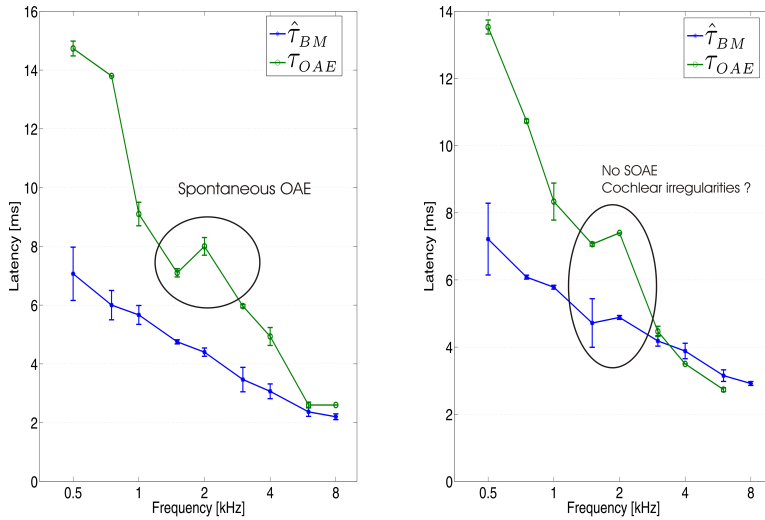


Figure 5.2: Comparison of OAE latency ( $\tau_{OAE}$ , green curve) and BM latency estimate ( $\hat{\tau}_{BM}^{(ABR)}$ , blue curve) for two illustrative subjects. Subject 9 on the left presents a "knee" around 2 kHz for the OAE curve and this knee is absent in the ABR measurement. A time-frequency analysis showed that SOAEs could account for this behaviour. On the contrary, subject 15, presented right, did not present SOAE activity and the plateau seen around 1-2 kHz may or may not be present in the ABR recording. The intra-subject variability is indicated by the vertical lines ( $\pm 1$  std).

### 5.2.1 Intra-subject variability

The intra-subject variability of both ABR and OAE measurement were shown in figures 4.5 and 3.9 respectively; their values are reported in appendix D. The trend can be more easily seen in Fig. 5.2 for two subjects. For both measurement methods, the standard deviation of each of the eleven subjects is rather small, of the order of 0.8 ms or less for OAEs and slightly more for ABRs (std < 1.13 ms). This good reproducibility of the data indicates that both techniques are reliable. However, at low frequencies, ABRs appear to be more variable than OAEs. This is due to the difficulty of detecting wave V compared to detecting the OAE burst. A reason for this more difficult detection could be the lower signal strength for ABR compared to OAE. This could be due to a decrease of neural synchrony at low frequencies which can be a result of the longer stimulus rise time used at these frequencies.

### 5.2.2 Inter-subject variability

Figure 5.3 presents the mean latencies calculated across all eleven subjects for both OAE and ABR measurements. The green line represents the OAE latency,  $\tau_{OAE}$ , and the blue line shows  $\hat{\tau}_{BM}^{(ABR)}$ . The inter-subject variability alone does not give an indication of the importance of the variability. It is necessary to consider the actual values of the latency estimates at each frequency. Table 5.1 shows the relative error of the inter-subject variability, i. e. the standard deviation relative to the latency value.

Frequencies [kHz]	0.5	0.75	1	1.5	2	3	4	6	8
$\tau_{OAE}$ variability [%]	7.1	11.8	14.8	10.9	8.1	10.9	12.1	18.2	10.4
$\hat{\tau}_{BM}^{(ABR)}$ variability [%]	9.1	9.9	18.3	12.9	15.3	17.9	17.5	19.2	20.1

Table 5.1: Inter-subject variability for  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$ , in percentage, across frequency. The percentage is calculated as the ratio between the standard deviation and the mean value.

It appears that the inter-subject variability of  $\hat{\tau}_{BM}^{(ABR)}$  is greater than the  $\tau_{OAE}$  inter-subject variability for almost all frequencies (eight out of nine). This means that the latency estimates differ more between subjects for ABR than for OAE. At cochlear level, subjects can indeed present different inhomogeneities along their basilar membrane, variably affecting the travelling wave. They may also have



distinctive cochlear filtering properties as well as a different threshold across the audible frequency range (Don et al., 1994). These differences affect the OAE and ABR inter-subject variability. On top of that, variabilities also occur at a neuronal level, due to different head sizes or gender (see chapter 2). This may explain why the inter-subject variability is greater for  $\hat{\tau}_{BM}^{(ABR)}$  than for  $\tau_{OAE}$ .

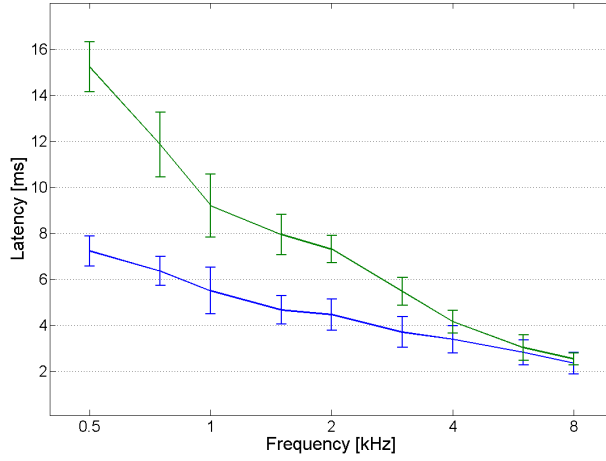


Figure 5.3: Average of OAE and ABR data for the eleven subjects. The green curve represents the estimate of  $\tau_{OAE}$  and the blue line  $\hat{\tau}_{BM}^{(ABR)}$ . The vertical lines represent the inter-subject variability ( $\pm 1$  std).

### 5.2.3 Acoustic wave hypothesis - $\tau_{OAE} = \hat{\tau}_{BM}^{(ABR)}$

A comparison between  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  can be seen in figure 5.2. For these two subjects, the curve representing  $\hat{\tau}_{BM}^{(ABR)}$  is well below the  $\tau_{OAE}$  curve. However, as frequency increases, the two curves get closer. This trend is the same for all subjects, as shown in appendix F. The average across subjects (figure 5.3) also confirms this behaviour.

The consequence is that the relation  $\tau_{OAE} = \hat{\tau}_{BM}^{(ABR)}$  might be true at high frequencies. However, this relation does not seem to hold at lower frequencies. It is not surprising to find  $\tau_{OAE} = \hat{\tau}_{BM}^{(ABR)}$  at high frequencies since the high frequency region

on the BM and the oval window are very close to each other; the travel time inside the cochlea for these frequencies is hence extremely short. But for the low-frequency region, it does not seem correct to neglect the backward travelling time.

#### 5.2.4 CRF theory - $\tau_{OAE} = 2\hat{\tau}_{BM}^{(ABR)}$

If the delay-frequency relation is represented by a power law, it appears as a straight line on a loglog axis. This makes the visual comparison of  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  easier. Linear functions were fitted to the data with variables  $a$  and  $b$  minimizing the distance between the experimental curve  $\tau_{OAE}$  and the theoretical curve  $bf^{-a}$ , where  $f$  is the frequency. The solid lines in Fig. 5.4 show the best fit for the ABR data (blue) and the OAE data (green), respectively. Besides, the slopes are only used for visually assessing the similarity between  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$ . If the CRF theory is exact,  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  should be on top of each other. If they are parallel to each other, this would support a factor different from 2 since, on a loglog axis, a multiplication factor just shifts the curves. To consolidate this visual analysis, a two-way analysis of variance (ANOVA) was carried out. It examines the effect of independent factors on the BM latency estimate. The independent factors are the frequency ( $n = 9$ ) and the measurement technique ( $n = 2$ , ABR or OAE). The null hypothesis is: The estimate of  $\tau_{BM}$  does not differ between techniques. Results are declared significant if the p-value is less than 0.05 and this would cast doubt on the null hypothesis. This ANOVA test is also a way to compare the slope of each curves.

The results of the ANOVA test for two exemplary subjects are presented in figure 5.4. The value of  $p$ , resulting from the two-way ANOVA test is indicated on the plot. For subject 10 (Fig. 5.4(a)),  $p = 0.0085$ , since  $p < 0.05$ , OAE and ABR measures are significantly different. In other words, the two techniques used do not estimate the same rate of change of latency, i.e.  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  are statistically different. For subject 6 (Fig. 5.4(b)), the two fitting curves are parallel to one another and  $p = 0.1446$ . The conclusion for subject 6 is that ABR and OAE describe the BM latency as a function of frequency in the same way.

The ANOVA test was run on the eleven subjects. For nine of them,  $p < 0.05$ . This means that  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  are significantly different in most cases (see appendix G for the results in all subjects). The first implication is that the CRF theory cannot

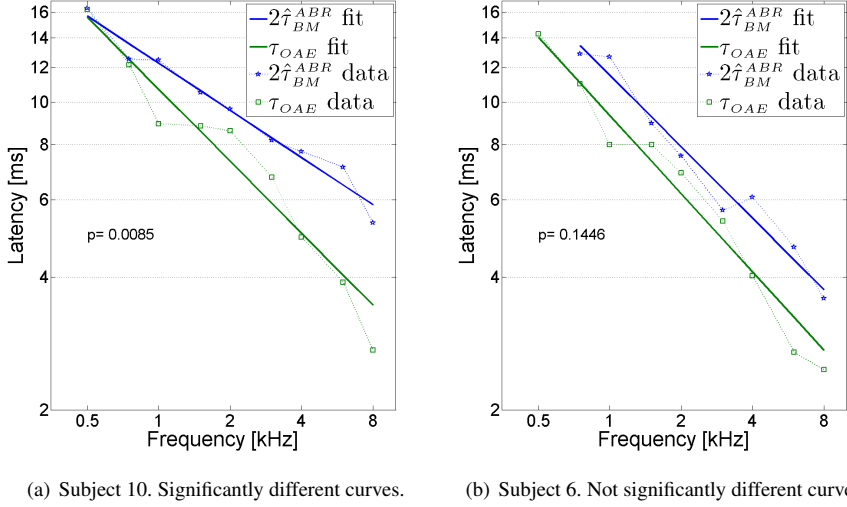


Figure 5.4: Comparison between  $\tau_{OAE}$  (green curve) and  $2\hat{\tau}_{BM}^{(ABR)}$  (blue curve) for two subjects. The data points are also plotted (symbol  $\star$ ) and are connected by dotted lines. The two solid lines represent the best fit to the data. The value of  $p$ , resulting from the two-way ANOVA test is indicated. For subject 10 (a),  $p < 0.05$ , it is therefore considered that OAE and ABR are significantly different but for subject 6 (b),  $p > 0.05$ , implying that the estimates given by OAE and ABR measurements evolve in the same way.

be supported by these results since the data obtained in this study do not verify  $\tau_{OAE} = 2\hat{\tau}_{BM}^{(ABR)}$  for most of the subjects. From this observation, different postulates emerge:

1. ABRs and OAEs are not generated at the same place. It could be that ABRs and OAEs represent two different mappings of the cochlear tonotopic organization (see Fig. 5.5). In such a case, a small variation of the ABR generation site would slightly change the value of  $2\hat{\tau}_{BM}^{(ABR)}$  and would align the two curves, provided this change is systematic with frequency. There is, however, a limit to the amount by which  $2\hat{\tau}_{BM}^{(ABR)}$  can vary, that is the distance between the OAE and ABR generation sites. The current state of knowledge indicates that they should be located where the BM is sufficiently excited, otherwise not enough hair cells are activated. Based on the spectra of the tone bursts (appendix A), it

seems acceptable to assume that one octave above the centre frequency, the BM does not vibrate enough to activate hair cells. To verify the hypothesis about a possible difference in generation site locus, the curve  $2\hat{\tau}_{BM}^{(ABR)}$  is shifted from its original position. ANOVA tests are run for  $2\hat{\tau}_{BM}^{(ABR)} \pm \delta$ , where  $\delta$  is varied between 0.1 and 5 ms. These boundaries are extreme values, knowing that a variation of 0.1 ms is insignificant compared to the values of  $2\hat{\tau}_{BM}^{(ABR)}$  and an error of 5 ms is not realistic, since this would mean that OAE and ABR are generated at completely different regions of the cochlea. An example is shown for subject 10 in figure 5.6. The dashed red lines limit the regions where the statistical significance would change; between these red lines,  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  become non-significantly different. For this subject 10, the  $\hat{\tau}_{BM}^{(ABR)}$  curve has to be shifted down by 1.4 ms to change the conclusion. For a 3-kHz tone burst, this would be more than an octave shift and it is therefore very unlikely. Similar observations were made for the ten other subjects. The conclusion about this postulate is that even with a small variation in the value of  $2\hat{\tau}_{BM}^{(ABR)}$ , the two curves still have significantly different slopes. This postulate can therefore be rejected.

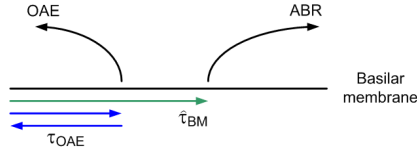


Figure 5.5: Postulate 1: ABR and OAE are not generated at the same place, it could be that ABR and OAE represent two different mappings of the cochlear tonotopic organization.

2. A second postulate is that OAEs do not travel in the same way back as the inward travelling wave that created them, i. e. there is no exact round-trip and OAEs might travel faster on their way back. This suggests that the factor 2 might not be the correct one. For instance, Shera and Guinan (2003) found factors relating  $\tau_{BM}$  to  $\tau_{OAE}$  of 1.7 and 1.6 for cats and guinea pigs, respectively. But then again, whatever the factor is, this reflects just a shift between the curves,  $\log(n\hat{\tau}_{BM}^{(ABR)}) = \log(n) + \log(\hat{\tau}_{BM}^{(ABR)})$ . The slope of the curves would not be affected and this second postulate can therefore be rejected.

3. A third postulate would be to restrict the relation  $\tau_{OAE} = 2\hat{\tau}_{BM}^{(ABR)}$  to a certain frequency range. A visual inspection of the data points in Fig. 5.4 (left) suggests that the slopes of the curves differ only if low frequencies (below, say, 2 kHz) or only high frequencies are considered. This postulate can be verified by "breaking" the curves at a certain frequency and running the same statistical analysis. This possibility will be discussed in section 5.3.

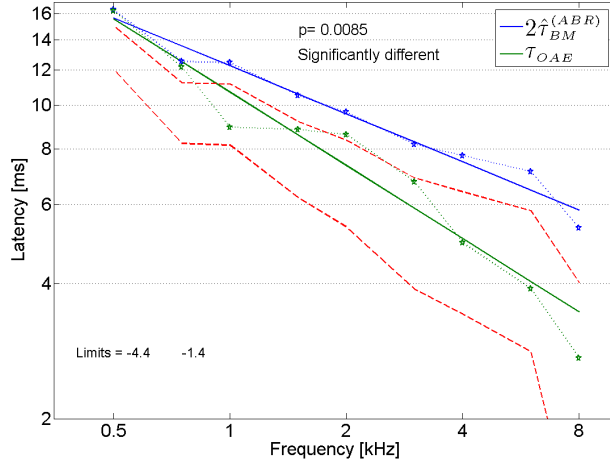


Figure 5.6: The curve representing  $2\hat{\tau}_{BM}^{(ABR)}$  was shifted from its original position for subject 10. The dotted red lines limit the regions where the significance would change, between these two lines OAE and ABR become non-significantly different. The boundaries are here 1.4 ms and 4.4 ms. A slight deviation from the values of  $\tau_{neural}$  and  $\tau_{synaptic}$  will not produce a change in the relation between  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$ . This observation is similar for all subjects and rejects the first postulate.

### 5.2.5 Signal-front hypothesis - $\tau_{OAE} = \tau_{front} + \hat{\tau}_{BM}^{(ABR)}$

An estimate of  $\tau_{front}$  obtained from human cadavers shows that it is frequency dependent and it becomes negligible above about 2 kHz (Ruggero and Temchin, 2007, Fig.7; Ruggero, 2007 and Tab. 5.2). For high frequencies, the signal-front hypothesis becomes identical to the acoustic wave hypothesis, i. e.  $\tau_{OAE} = \hat{\tau}_{BM}^{(ABR)}$ .

Frequencies [kHz]	0.5	0.75	1	1.5	2	3	4	6	8
$\tau_{front} \cdot 10^{-2}$ [ms]	71.46	38.11	20.33	5.78	1.65	0.13	0.01	0.00	0.00

Table 5.2: BM front delays for different frequencies. These values were obtained by fitting a curve to the experimental data given by Ruggero since he provided delays between 0.1 and 1.4 kHz only. The original data from Mario Ruggero were taken from an experiment that von Békésy conducted on human cadavers and were compensated for the effect of death.

The comparison between  $\tau_{OAE}$  and  $\tau_{front} + \hat{\tau}_{BM}^{(ABR)}$  is shown in figure 5.7 for one illustrative subject. The plots for all subjects can be found in appendix H.

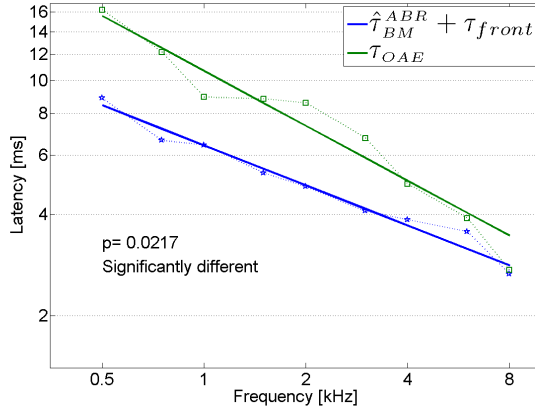


Figure 5.7: Comparison between  $\tau_{OAE}$  (green curves) and  $\tau_{front} + \hat{\tau}_{BM}^{(ABR)}$  (blue curves) for subject 10.  $\tau_{front}$  has been obtained from Ruggero (2007). It should be noted that the curves are very close to each other at high frequencies but further apart for low frequencies.  $p$  is the result of the ANOVA test, in this example  $p < 0.05$  meaning that there is a significant difference between  $\tau_{OAE}$  and  $\tau_{front} + \hat{\tau}_{BM}^{(ABR)}$ .

An ANOVA test was run in order to estimate the deviation between the two values. The results show that six out of eleven subjects present a significant difference between  $\tau_{OAE}$  and  $\tau_{front} + \hat{\tau}_{BM}^{(ABR)}$  (refer to appendix H for the complete results). The results are not as clear as for the CRF theory. It is therefore a bit more difficult to decide whether the signal-front theory offers a good estimate of  $\tau_{OAE}$  or not. One key difference with the other theories is here the prediction of a dissimilarity between high and low frequencies, which is in agreement with the data. The signal-front hypothesis could be slightly modified to comply with the present data. It might be

that  $\tau_{front}$  is underestimated at low frequencies. Another possibility is that the choice of tone-bursts rise time has an influence on wave-V latency and hence on  $\hat{\tau}_{BM}^{(ABR)}$ . This hypothesis was raised by Ruggero and Temchin (2007) who claimed that the difference in stimulus rise time in Neely et al. (1988) added a supplementary delay to  $\tau_{wave\ V}$ . They argued that identical rise times are necessary to have the synchronous neural firing occurring at the same time for all the stimuli. Since the tone bursts of the present study also have different rise times, the effect of the rise time on the wave-V latency was investigated and is presented in appendix E. It is concluded that having the same rise time for all frequencies would not significantly change the relative position of the  $\hat{\tau}_{BM}^{(ABR)}$  and  $\tau_{OAE}$  curves.

The conclusion about the signal-front theory is that it does not hold for most of the subjects over the entire frequency range. However, for 4-8 kHz, the present data confirms the prediction  $\tau_{OAE} = \tau_{front} + \hat{\tau}_{BM}^{(ABR)}$  and a slight change in the value of  $\tau_{front}$  could make this theory comply with the data of this study for the entire frequency range. Compensating for the rise time did not bring a more accurate explanation about the difference that appears between high and low frequencies. The previous section about the CRF theory concluded that the CRF theory could be investigated further if high and low frequencies were analysed separately, this is discussed in the next section.

### 5.3 Discussion

Three current theories for OAE generation were presented in the previous section and their prediction for  $\tau_{OAE}$  confronted with the present results. The conclusion drawn for the acoustic wave and signal-front hypotheses is that they hold only for limited frequency range ( $f \gtrsim 4$  kHz). For the CRF theory, results seem to give a different conclusion whether the whole frequency range was considered or only part of it. Further investigations are therefore needed to draw a clearer conclusion about this theory. The following section presents a more detailed analysis of the CRF hypothesis when high and low frequencies are studied separately.

### 5.3.1 Separating the analysis below and above 2 kHz

As the third postulate of section 5.2.4 suggested, a further analysis of the data was carried out, in which high and low frequencies were treated separately. The idea is to run a statistical analysis to compare  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  for each frequency range. The break point was chosen to be at 2 kHz. This enables also a direct comparison with the study of Neely et al. (1988) who compared OAE and ABR data up to 2 kHz. It was found that ten out of eleven subjects did not show any statistical difference between  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  at low frequencies ( $f < 2$  kHz, ANOVA  $p > 0.05$ ) and six out of eleven at higher frequencies. An example is illustrated in figure 5.8.

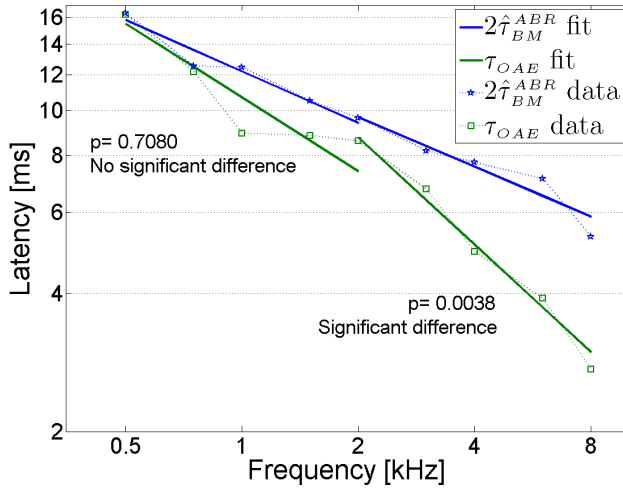


Figure 5.8: Latency estimates from OAE ( $\tau_{OAE}$ , green curves) and ABR ( $2\hat{\tau}_{BM}^{(ABR)}$ , blue curves) for subject 10. The dotted lines represented the data and the solid lines are a fit to these data. The results of the ANOVA test are indicated above the curves. The behaviour shown here is identical for the majority of the subjects: at low frequencies ( $f < 2$  kHz) there is no significant difference between  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  ( $p > 0.05$ ) but there is a significant difference at higher frequency ( $p < 0.05$  for  $f > 2$  kHz).

The direct consequence is that there appears to be a difference in behaviour between low and high frequencies. The relation suggested by the CRF theory ( $\tau_{OAE} = 2\hat{\tau}_{BM}^{(ABR)}$ ) seems to hold at low frequencies ( $f < 2$  kHz) but not at higher ones ( $f > 2$  kHz). The conclusion for the frequency range 0.5-2 kHz is consistent



with the data from Neely et al. (1988). They compared  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  between 0.5 and 2 kHz, as an average across their subjects (7 for OAE and 20 for ABR). This comparison is replotted in figure 5.9. The OAE data of the present study were averaged (green curve with standard deviation) and compared with a function fitted to the ABR data (blue line). Note that the data from Neely et al. (1988) were only available up to a level of 55 dB SPL, while a level of 66 dB pe SPL was used in the current data. This lower level might partially explain the higher latencies in Neely et al. (1988). The study of Neely et al. concluded that the comparison between OAE and ABR data “shows agreement between these two measures that is consistent with "cochlear echo" theory for the origin of the OAE”. The "cochlear echo" theory is similar to the more recent CRF theory and predicts the same round-trip for the OAE latency, i. e.  $\tau_{OAE} = 2\tau_{BM}$ . As shown in figure 5.9 (right), the ABR fitting curve of the present study suits the OAE data nicely. The data from the present study cast however doubt on the possibility to extend the prediction  $\tau_{OAE} = 2\hat{\tau}_{BM}^{(ABR)}$  to frequencies above 2 kHz.

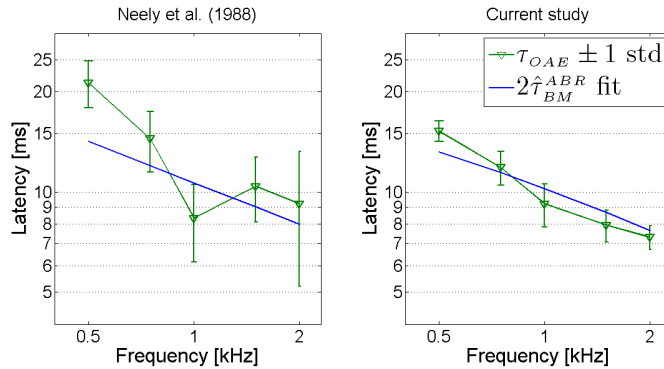


Figure 5.9: Comparison between the study from Neely et al. (1988) and the present study. The symbols represent the averaged OAE latency across subjects, the error bars indicate  $\pm 1$  standard deviation and the straight lines is the fit to  $2\hat{\tau}_{BM}^{(ABR)}$ . With the present data, ABR recordings follow the OAE data very well.

There seems to be a different behaviour between low and high frequencies and this was also observed in previous studies comparing ABR and OAE in humans (Narayan, 1991) or directly measuring BM motion in animals (Shera and Guinan, 2003; Siegel

et al., 2005). It is still unclear where this difference originates. Shera and Guinan (2003) suggested that this could be due to a breakdown of the scaling symmetry whilst Siegel et al. (2005) suggested theoretical implications for OAE generation, as described below.

### 5.3.2 Scaling symmetry

The BM can be seen as a series of filters, each centred at a different frequency, CF. Each of these filters has a different bandwidth, which is CF-dependent. The scaling symmetry is a property of the BM which assumes that the  $Q$  factor<sup>1</sup> of the BM filters, i. e. the reciprocal of their relative bandwidth, is constant across frequencies. In other words, when the frequency doubles, the bandwidth of the filter doubles, too. Filters are also characterised by their impulse response, where broad filters have short impulse responses and vice versa. Figure 5.10 (left) illustrates how this impulse response length,  $\tau$ , evolves with CF. When both  $\tau$  and CF are plotted on a log axis, their relation appears as a straight line (due to the cochlear tonotopic mapping). This is an ideal case where the scaling symmetry holds.

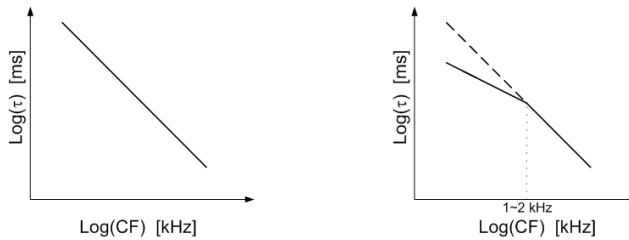


Figure 5.10: BM filters impulse response length,  $\tau$ , as a function of centre frequency, CF. On a loglog axis their relation appears as a straight line. Left: Ideal cochlea where the scaling symmetry holds. Right: Actual cochlea with a breakdown of the scaling symmetry below 1-2 kHz.

Experiments have indeed shown that below about 1 kHz, the  $Q$  factor decreases (Fig. 3.5 in Moore, 2003). The consequence is a relative broadening of the filters below about 1 kHz compared to the "ideal" case, which also implies that their impulse response becomes shorter. This is illustrated in Fig. 5.10 (right). The relation

<sup>1</sup>  $Q = \frac{CF}{\Delta_f}$ , where  $\Delta_f$  is the bandwidth of the filter centred at CF.

between  $\tau$  and CF is no more a straight line across the frequency range. This figure is very similar to Fig. 5.8, where the low frequency part of  $\tau_{OAE}$  has a shallower slope than the high frequency part. The present results thus seem to confirm previous studies showing that the scaling symmetry is not holding across the whole frequency range.

One reason for this "breakdown" of the scaling symmetry could be the difference in activity of the cochlea amplifier. Experimental observations showed that "at the extreme cochlear apex, positive feedback (the cochlear amplifier) plays a lesser role than at the base" (Robles and Ruggero, 2001). This change between high and low frequencies results in a change in the excitation pattern shape along the basilar membrane, as illustrated in Fig. 5.11. At high frequencies, the filters become sharply tuned to a specific frequency while, at low frequencies, the filters are broader and embrace a wider range of frequencies. The consequence on OAE generation is that if the BM is excited over a larger area, it is more likely that the OAE generators (wherever they precisely are) will cancel each other's contribution to the backward travelling wave. The consequence of such cancellations would be a less detectable waveform. And a waveform rising from a more basal area (lower latency) might be wrongly assigned to the OAE burst. This might explain what is seen in Fig. 5.8.

Regarding the ABR data (blue curve in Fig. 5.8), things are not so clear. ABR do not seem to be affected in the same way by this difference in cochlear activity between high and low frequencies. The difference between the OAE and the ABR curves is not constant across frequencies and is not only a shift. It seems that OAE and ABR reflect two different mappings of the cochlea. Further investigations are required to fully understand these differences.

In a very recent paper, Moleti and Sisto (2008) also compared ABR and OAE latencies across a wide frequency range. Their conclusion is very similar to the ones presented here. They estimated  $Q$  (their Fig. 6) and found out that its value stays constant down to about 2.5 kHz and then decreases. These similar results encourage further investigations on where and how the difference between ABR and OAE generation might occur.

To conclude on the hypothesis suggested by the CRF theory, it seems that the prediction  $\tau_{OAE} = 2\hat{\tau}_{BM}^{(ABR)}$  holds for frequencies below 2 kHz but not at higher

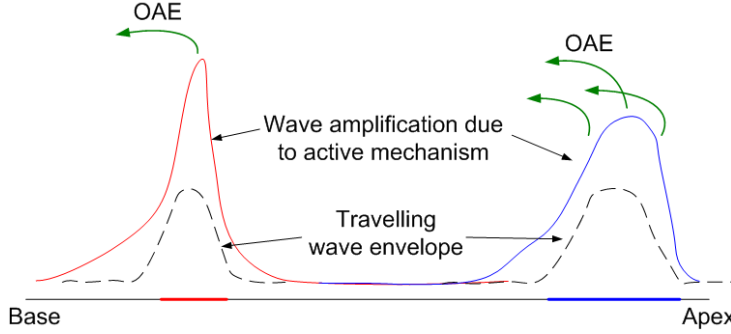


Figure 5.11: This sketch shows the effect of the active mechanism on an incoming wave. The stimulus is a tone burst, exciting a narrow region of the basilar membrane. The high frequencies are processed at the base of the cochlea and the low frequencies at the apex. As it was shown in previous studies (Robles and Ruggero, 2001), the active mechanism has a bigger influence at high frequencies. The BM presents a gradient stiffness and can be seen as a series of small oscillators uncoupled with one another. At the high frequencies these oscillators are in phase because a narrow region is excited, the neighbouring regions will create out of phase waves of much smaller amplitude. The backward travelling wave from the CF region is therefore detectable. When scaling symmetry breaks, the cochlear amplifier is not so active and a broader region of oscillators will react. In this case, there is the possibility for wave cancellation, the backward travelling wave generated at CF might not be dominant.

frequencies. This difference of behaviour could be explained to some extent by the discontinuity of the scaling symmetry of the basilar membrane.

## 5.4 Conclusion

The data obtained in the OAE and ABR experiments showed good reproducibility for each individual subject. It can therefore be concluded that the results obtained in this study are based on reliable and robust measurements. The outcome of the present study concerns two main points: a) The generation site of OAEs and ABRs and b) the way OAEs travel back to the ear canal.

This study suggests that the CRF theory holds at low frequencies,  $f \lesssim 2$  kHz. This reinforces the prediction for the OAE delay being twice the delay of the forward travelling wave,  $\tau_{OAE} \approx 2\hat{\tau}_{BM}^{(ABR)}$ , below 2 kHz only. The present results are in contradiction with previous studies that showed a deviation from  $\tau_{OAE} \approx 2\tau_{BM}$

at low frequencies (Shera and Guinan, 2003; Siegel et al., 2005). The differences in the experiments could, to some extent, explain this discrepancy. The physiological data were collected from animals (cats and guinea pigs for Shera and Guinan (2003), chinchillas for Siegel et al. (2005)) comparing stimulus-frequency OAEs with BM vibration or auditory-nerve fibres responses. These techniques require surgical invasion that can lead to cochlear fluid leakage and dramatic changes of the BM properties. Despite the care taken by the studies, the state of the examined cochleae might still have been altered. The recordings of TBABR and TBOAE on humans are non-invasive but also more indirect. This is a trade-off to make to avoid modifying the state of the cochlea. The signal-front hypothesis appears to hold for high frequencies ( $f > 4$  kHz) confirming the idea of a fast backward travelling wave. The theoretical implications of these findings refer to the generation site of OAEs and ABRs which might not be the same along the BM partition. At high frequencies the present data support the fast wave hypothesis, it seems that shortly after the stapes enters in vibration to produce the inward travelling wave in the cochlea, a compressional wave is created at the CF place and travels backward through the cochlear fluid. This fast reaction is also linked with the cochlear amplifier which is more active at higher frequencies. This difference in cochlear activity between high and low frequencies could explain the different trends observed for  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  (figure 5.8). At low frequencies ( $f \lesssim 2$  kHz), the distance between the mean of  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  is minimum for  $\tau_{OAE} = 1.9 \hat{\tau}_{BM}^{(ABR)}$ . There might therefore be other mechanisms involved for the OAE to go back to the ear canal. The fast-wave hypothesis can be discarded for the low frequencies and, as Mountain confirmed in Gummer (2003, page 585), "There is more than one mode (of propagation) in the cochlea but in terms of the low-frequency, it seems that both forward and reverse travelling waves are propagating by the same modes". However, the present data do not contradict the possibility of a backward travelling wave through the tectorial membrane instead of the BM. Using a scanning laser interferometer, Ren (2004) measured the BM vibration in Gerbils and did not observe any backward travelling waves.

The possibility of a multiple mode travel is not excluded either. The work from Narayan (1991) suggests that the backward travelling wave uses two different modes, one through the cochlear fluid and one along the cochlear partition, explaining

the two bursts in her OAE recordings. It could be that, at low frequencies, the wave travelling through the fluid is cancelled or reflected by irregularities on the BM and therefore does not reach the stapes. The other (slow) wave created around the CF place uses either the BM or other membranes to travel back to the base. For high frequencies, however, the more active cochlear amplifier may produce a stronger backward wave which might not be as disturbed by the irregularities on BM or the interactions with other waves. It could be said that the fast wave is the dominant propagation mode at high frequencies.

The possible explanation for the factor of 1.9 between  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$ , observed at low frequencies, could be that the resulting travelling wave is generated at regions basal to the peak of the BM. As seen in figure 5.11, the relative BM filter bandwidths are broader at low frequencies. Thus, a larger region of the BM is excited resulting in a multitude of backward travelling waves. Due to the spread region, there is more chance that the waves are out of phase and the oscillators from sites above CF might cancel the wave created at CF. But the waves created basal to CF will not undergo such cancellation, resulting in an earlier wave. This idea of short latency contribution and long latency cancellation has also been mentioned in Siegel et al. (2005).

The scaling symmetry was found here to break down at some point along the BM, but where should the distinction between high and low frequencies exactly be? The present study suggests a limit at 2 kHz and Siegel et al. (2005) observed a limit at 3-4 kHz in chinchillas. More measures of the BM vibration in humans would be needed to know more precisely where the transition occurs.

With the present data it could be interesting to study further the different models available for the cochlea (de Boer et al., 2007; Shera et al., 2007). The findings about the mode of wave propagation could be compared with models that simulate OAEs generation. Finally, the effect of the middle ear on OAEs could also be included. These aspects were beyond the scope of this study but remain a source of possible future projects.

## Individual steady-state response delays

---

### Abstract

The previous chapters showed that an estimate of the cochlear delay can be obtained from ABR and OAE recordings. There exist other methods to estimate auditory delays such as auditory steady-state responses. ASSRs evoked by AM tones modulated in the 80-Hz range are generated in the brainstem and are therefore comparable in terms of latency to the ABR. Such comparison has not yet been published for normal-hearing adults. This experiment estimated ASSRs latencies in 13 normal-hearing subjects using AM tones at six different carrier frequencies. The results show good reproducibility of the measures for each subject and a good agreement with the ASSR latencies published in previous studies. A difference of about 1 ms is found between the ASSR latencies of male and female subjects. The ASSR latencies are about 10 ms longer than the TBABR delays, it is hypothesized that 80-Hz AM tones evoke responses at higher stages of the brainstem, probably in the inferior colliculus and above. Both ABRs and ASSRs present the same decrease of latency as a function of frequency.

### 6.1 Introduction

The results of the ABR and OAE experiments were compared and new evidence is found about their generation sites and about the OAE travel mode. However, the ABR experiment showed some difficulties in analysing the responses at low frequencies, where the detection of wave V was sometimes difficult because of a too high noise

level. As seen in Chap. 2, auditory steady-state responses can measure the activity of the brainstem and the peak obtained in the frequency domain for an ASSR is usually sharper and higher than the wave V from a TBABR. This makes an ASSR a more detectable signal than a TBABR. Moreover, stimuli used to record an ASSR are AM-tones which have a higher frequency specificity than the tone bursts used to record an ABR. This means that an ASSR gives information about a limited region of the cochlea, which is an important advantage when one wants to estimate auditory delays. The previous two methods have estimated cochlear delays based on OAEs and ABRs, a relation between their delay,  $\tau_{OAE}$  and  $\tau_{wave\ V}$ , and the BM delay,  $\tau_{BM}$  could be found since the origin of both OAEs and ABRs is roughly known. On the contrary, the origin of ASSRs is rather approximate, depending on the modulation rate of their stimulus, ASSRs are generated between the brainstem and the auditory cortex. If one wants to record ASSRs evoked in the brainstem, the exact modulation frequency to use is not precisely known, previous studies only mentioned the use of a modulation rate above about 50 Hz (Sininger and Cone-Wesson, 2002; Kuwada et al., 2002). ASSRs can therefore not be used to estimate cochlear delays directly. That is why most studies have focused on hearing threshold estimation based on ASSRs rather than on their actual latency (Kuwada et al., 1986; Lins and Picton, 1995; Cone-Wesson et al., 2002). It is however assumed that the difference in latency between frequencies occurring at the cochlear level is reflected in the ASSRs, i.e. high-frequency stimuli have shorter delays than low-frequency stimuli. ASSRs can hence be used to assess the inter-frequency delay, which is the cochlear delay measured at a higher stage in the auditory pathway. Only few studies have estimated ASSR delays and even fewer tried to relate this delay with other auditory delays. John and Picton (2000) combined the ASSR inter-frequency delay and DPOAE delays to estimate  $\tau_{filter}$  and  $\tau_{transport}$  (see Chap. 2), they concluded that their analysis over simplifies the OAE generation mechanism and cannot give a precise estimate of these two delays. Purcell et al. (2003) also used DPOAE latency and estimated the delay between the BM and the ASSR generator at certain frequencies. They concluded that this inward conduction delay was incorrectly estimated due to the compromises they made in order to simultaneously record DPOAEs and ASSRs such as low stimulus levels and different frequencies. There is therefore a lack of comparative studies



about auditory delays. The present study proposes to compare delays of ABRs and ASSRs evoked with similar stimulus level. Since ABR delays were already compared with OAE delays in the previous chapter, the present chapter completes this unique comparison of delays from the auditory system. The present chapter describes the experiment run to obtain brainstem delays estimated from ASSRs in normal hearing subjects. First, the method used to derive latencies from ASSRs is presented. Then, similarly to the previous experiments, this chapter investigates the intra- and inter-subject variabilities. So far, no study has been published to directly compare ABR and ASSR delay estimates in normal hearing adults, this chapter therefore fills the lack of data comparison between ABRs and ASSRs. These two techniques are also compared in terms of their reliability and robustness. The latency of ABRs and ASSRs as a function of frequency present the same slope and it is shown that an ASSR has a higher latency than an ABR. Suggestions are made to explain this result.

## **6.2 Methods**

### **6.2.1 Subjects**

The subjects participating in this experiment were 13 normal-hearing adults: 7 females and 6 males, aged between 18 and 31 years. They all had pure-tone thresholds better than 15 dB HL in the range 0.25-8 kHz. This experiment was run at the Hearing and Balance Centre of Southampton University, it was therefore not possible to have the same subjects as in the ABR and OAE experiments, which were carried out at DTU. That is why only the mean latency values across all subjects will be compared. The subjects were seated in a comfortable chair in a sound insulated booth. They were asked to move as little as possible during the session which lasted about one hour. Some subjects reported falling asleep, which should not affect ASSR if they are generated in the brainstem (Chap. 2). The experiment was approved by the departmental Safety and Ethics Committee of the Institute of Sound and Vibration Research (ISVR).

### 6.2.2 Stimuli and recording procedure

The stimuli used to elicit steady-state responses are amplitude modulated tone bursts at six different CF: 0.5, 0.8, 1, 1.6, 4 and 6.3 kHz. The choice of these CF was influenced by the previous ABR experiment and by the duration of the experiment. Using six frequencies kept the experiment below an hour, which seems an acceptable duration for voluntary subjects. In order for all the signals to start with the same phase, it was necessary to fit an entire number of cycles inside the averaging period of 204.8 ms, some values of the CFs were then adjusted and the exact values used are 498, 800, 1001, 1602, 3999 and 6299 Hz. They will be referred to as their rounded values for simplicity. Three CFs were presented simultaneously, 0.5-1-4 kHz in one run and 0.8-1.6-6.3 kHz in a second run. Studies have shown that a simultaneous presentation of tones modulated at different frequencies does not alter the ASSR amplitude too much and allows a shorter recording time (Lins and Picton, 1995). These tones were modulated in amplitude with a depth of 100% and, in order to have an entire number of cycles inside each frequency modulation ( $f_m$ ) cycle, it was here also necessary to adjust their values. The exact  $f_m$  values were 83.008, 87.89 and 92.774 Hz and will be referred to as 83, 88 and 93 Hz, respectively. As described in Chap. 2, an 80-Hz ASSR originates from the brainstem and is therefore a good choice to compare with ABR recording. CFs of 0.5 and 0.8 kHz were modulated at 83 Hz, 1 and 1.6 kHz at 88 Hz and 4 and 6.3 kHz at 93 Hz. Using different modulation frequency has no significant effect on ASSRs amplitude (John and Picton, 2000, Fig. 6). This means that the frequency at which the tones are modulated will not change the detectability nor the latency of the response. Each stimulus was repeated 600 times. Both stimulus generation and response recording used an ISVR software. The setup is illustrated in Fig. 6.1. The stimuli were sent to a Cambridge Electronics Design (CED) $\mu$ 1401 mkI D/A converter connected to a KC50 Audiometer. The level was adjusted to 66 dB SPL on this audiometer. This level is chosen to compare with the ABR experiment. For higher levels such as 75 dB SPL, the AM-tones presented simultaneously interfere with each other and lower the response amplitude (John et al., 1998). The signal is then sent to the subject's ear canal via an ER-5A insert earphone. Electrodes are placed on the subjects head: on the high forehead (positive), on the low forehead (ground) and on the neck (reference), as shown in figure 6.1. The electrodes are connected to the CED

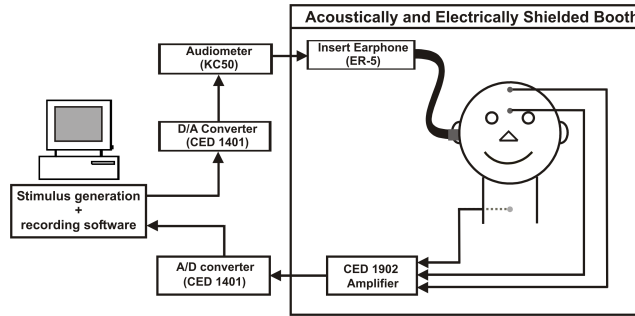


Figure 6.1: Equipment used for the recording of ASSR. The stimuli are produced by a computer, then sent to a D/A converter. The analog signal is transmitted to the insert earphone ER-5. The responses are recorded by an amplifier and sent to a D/A converter. The electrodes are positioned on the high and low forehead and on the neck.

1902 amplifier which sends the recorded signal to the CED  $\mu$ 1401 mkl A/D converter and then back to the computer where the responses are stored for off-line analysis. It is necessary to filter the signal before sending it to the A/D converter in order to remove some of the background noise. However, each time a filter is used, some phase shift appears. The phase of the response is a crucial parameter to estimate the latency. A preliminary study has investigated the optimal parameters to avoid noticeable phase lags. The responses were therefore low-pass filtered at 250 Hz and amplified by a gain of  $10^4$ . For the calibration, the stimuli were sent to the ER-5, plugged to an IEC 711 B&K ear simulator. The level was measured with the B&K 2231 sound level meter and its 1/3 octave filter. This filter allows separate measurements of the SPL at each of the frequencies contained in the AM tone-complex.

### 6.2.3 Response detection

The detection of the response at a given frequency uses the magnitude-squared coherence (MSC) method presented in section 2.2.2. Studies have shown that the MSC is a robust criterion when analysing ASSR in the 80-Hz modulation range (Schoonhoven et al., 2003). The MSC method returns a coefficient reflecting the robustness of the signal with regards to noise. The recorded signal was divided into 16 averages to calculate the MSC value since previous studies showed that, for a

number of averages  $q = 16$ , the 95% confidence limit is 0.183. If MSC is lower than this value, the recorded signal is very likely to be only noise (Dobie and Wilson, 1993) and the response is discarded.

#### 6.2.4 ASSR latency

The main purpose of this study was to estimate the ASSR latency at different frequencies. It is assumed that the ASSR latency measured is only frequency-dependent up to the cochlea, i. e. the synaptic delay and the neural conduction time are frequency independent (as described in Chap. 4 and Neely et al. (1988)). This can give valuable information about their generation site(s) since responses generated in the brainstem have shorter latencies than responses stemming from the cortex. For OAE measurements, the latency is estimated between the stimulus onset and the OAE peak; for ABR measurement, the latency is the time between the stimulus onset and the wave V peak. In these two cases, the detection is made in the time domain. Due to the steady-state nature of their stimuli, ASSR cannot be detected directly in the time domain and their latency can therefore not be estimated from the temporal response. However, the phase can be used to estimate the latency of the response. The phase delay,  $P$ , is calculated relative to the stimulus. This is done by subtracting the phase  $\theta$  of the response at the modulation frequency  $f_m$ , from the stimulus phase. In mathematical terms,  $P = 360 - \theta$  (Rodriguez et al., 1986). This phase delay can in turn be converted into latency by  $L = \frac{P}{360 \cdot f_m}$ . There is an ambiguity when using the phase of the signal as the number of stimulus cycles that occurred before the onset of the response is unknown. The calculated phase only indicates where on the waveform the response occurs and this can happen every  $2\pi$  due to the periodic nature of the signal. This ambiguity can be resolved by using the preceding cycle technique introduced by John and Picton (2000). This technique takes advantage of the fact that the latency of a response evoked by a tone modulated by frequencies in the same region (e. g. 80-Hz range) depends on its carrier frequency. An integral number of  $2\pi$  (or  $360^\circ$ ) is added to the phase delay  $P$  to estimate the latency,  $L$ , for different numbers of cycles. The

$f_m$ [Hz]	83	88	93	$\Delta t$ [ms]
$n = 0$	4.69	5.35	5.99	<b>1.3</b>
$n = 1$	16.74	16.73	16.77	<b>0.03</b>
$n = 2$	28.79	28.11	27.55	<b>1.24</b>

Table 6.1: ASSR latency for a 4 kHz tone as a function of modulation frequency for different number of cycles  $n$ . It appears that for  $n = 1$  the latency is constant around 16.7 ms, proving that the response occurs one cycle after the stimulus onset.

expression for the latency becomes

$$L = \frac{n360 + P}{360 \cdot f_m}, \quad n \in \mathbb{N}$$

The number  $n$  of preceding cycles was determined for one subject for different carrier frequencies (0.5, 1, 2, 4 and 6 kHz) modulated at 83, 88 and 93 Hz. This is illustrated in table 6.1. It appears that for  $n = 1$  the latency is constant, around 16.7 ms, across modulation frequencies, this observation is similar for the other carrier frequencies tested. This proves that the response occurs one cycle after the stimulus onset. This result is in agreement with experiments previously carried out at ISVR and also published studies (John and Picton, 2000; John et al., 2001). The number of preceding cycles stays the same throughout the experiment for any modulation frequencies considered.

The next section presents the results of the experiment run to estimate ASSR latency for different subjects.

## 6.3 Results

### 6.3.1 Latency-frequency functions

The latency estimates of the ASSR as a function of frequency for the 13 subjects are presented in figure 6.2. One of the observations is that latency decreases with increasing frequency. This result is expected and indicates that the ASSRs reflect the place-frequency mapping of the cochlea, i.e. high frequencies are processed before low frequencies and therefore have shorter delays. However, the decrease is

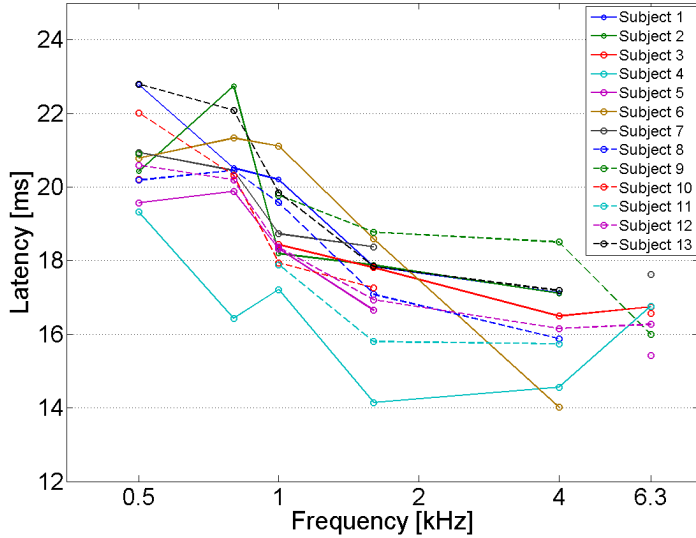


Figure 6.2: ASSR latency estimate as a function of carrier frequency for the 13 subjects. The general trend is a decreasing latency with increasing frequency.

not monotonic for some subjects, for instance between 0.5 and 0.8 kHz the latency increases (subjects 2, 5, 6, 8, 9), and similarly between 4 and 6.3 kHz (subjects 3, 4, 5, 12). To explain this phenomenon, a possibility is to investigate the intra-subject variability. If it is large at those frequencies, the latency increase observed reflects an inherent variability rather than an actual increase of the latency. The intra-subject variability will be studied in the next section. Another possibility is that there is some wrapping issue with the phase at some frequencies, as reported in previous studies (John and Picton, 2000). In that case,  $360^\circ$  need to be added or subtracted to the calculated phase. For example, subject 5 showed a latency of 19.6 ms at 0.5 kHz; if  $360^\circ$  was added to the corresponding phase the latency became 31.6 ms. This is much larger than the previous value and is not realistic compared to the latencies at other frequencies. The same conclusion applies for the other subjects at 0.5 and 6.3 kHz. For this latter frequency, there is another remark to make: the latency could

be estimated for only seven out of the 13 subjects. There appeared to be a high noise level for this frequency. This is in contradiction with Cohen et al. (1991), who showed that for modulation frequencies above 70 Hz, the response is more easily detectable. After investigating the equipment for possible noise sources, the precise reasons for the high noise level at 6.3 kHz remain unclear. Figure 6.2 also illustrates the case of subject 4, who is well below the other subjects. This outlier will be discussed more in detail in the section 6.4.2.

### 6.3.2 Intra-subject variability

In order to investigate the intra-subject variability, the experiment was repeated for eight subjects. Subject unavailability prevented from recording more subjects. The mean values and the standard deviation<sup>1</sup> of these eight subjects is presented in figure 6.3, their exact values are reported in appendix I. It was not always possible to obtain two values for each frequencies due to a high noise level. In such a case, the standard deviation is zero, e. g. for subject 2 at 0.8 kHz. The ASSRs showed a good repeatability, the highest intra-subject variability is 1.9 ms at 6.3 kHz for subject 5 and the lowest is 0 (perfect repeatability) for subject 13 at 1.6 kHz. These values are small compared to the latency values (approx. 16 ms). This makes ASSR a reliable method to measure individual ASSR latencies.

### 6.3.3 Comparison with previous studies

The comparison between the results obtained in the present study and previously published data is presented in figure 6.4. All these studies considered modulation frequencies between 70 and 100 Hz which elicit responses mainly from the brainstem. The standard deviations of the other studies are not shown for clarity. The first observation is that the latency estimates from the present study are generally in good agreement with results from previous studies. This suggests that the phase was correctly calculated, i. e. the unwrapping does not seem to have been a problem. The

---

<sup>1</sup> The standard deviation calculated here differs from the one calculated in the previous experiments, this is due to the lower number of repetitions. A biased standard deviation is calculated by dividing by the number of repetitions (here two), i. e.  $std = (\frac{1}{2} \sum_{i=1}^2 (x_i - \bar{x})^2)^{\frac{1}{2}}$ .

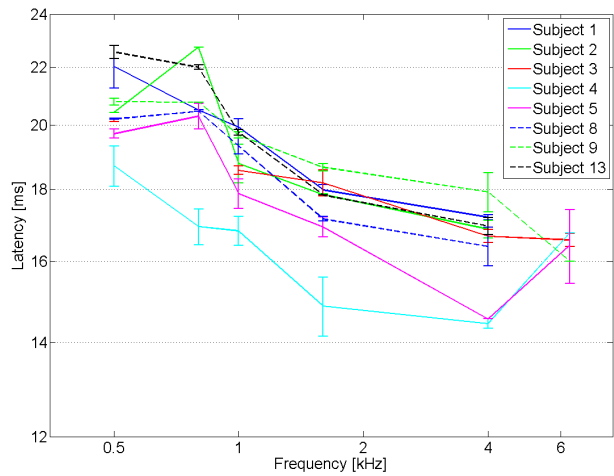


Figure 6.3: Intra-subject variability for eight subjects. The error bar indicates the latency difference between the two measurements.

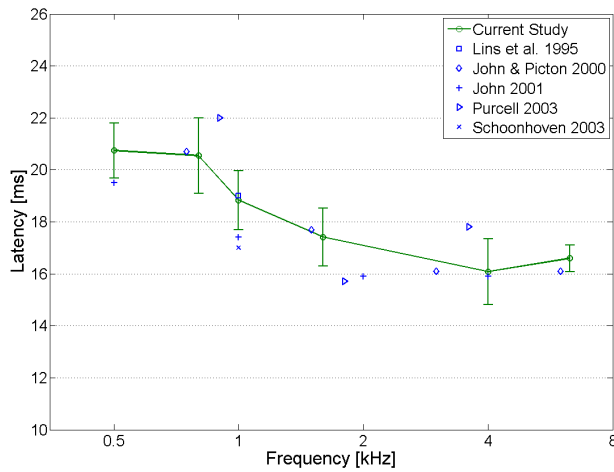


Figure 6.4: Comparison of the current results with previous ASSR studies. The error bars indicate the inter-subject variability ( $\pm 1$  std) and it shows that the results obtained here follow previously published results.



difference observed between studies can be explained by a different stimulus level; higher stimulus levels lead to shorter ASSR latency (see Chap. 2). For instance, the value from Schoonhoven et al. (2003) at 1 kHz is below the value estimated here and this can be explained by their higher level (70 dB nHL) and their great standard deviation ( $\pm 6$  ms). The results from John et al. (2001) are slightly below the current values, they do not mention their standard deviation which would give a better picture of the variability of their data. An observation made earlier is the increase of the latency, in the current study, between 4 and 6.3 kHz. This is not completely inconsistent with the other studies either. John et al. (2001) and John and Picton (2000) have constant latencies above 2 kHz and the latency estimate from Purcell et al. (2003) also increase between 1.8 and 3.6 kHz. The exact reason for such behaviour was not commented by John et al. (2001) nor John and Picton (2000) but Purcell et al. (2003) explained that the nerve fibres connecting the cochlea to the brainstem are shorter for the mid-frequencies, inducing a shorter delay at these frequencies. Further suggestions will be made when comparing with ABR. Figure 6.4 also illustrates the inter-subject variability across frequencies. It can be concluded that there is no effect of the frequency on the variability. The smaller value at 6.3 kHz is an effect of the lower number of values obtained at this frequency where data was available for only seven out of the 13 subjects (see appendix I).

## 6.4 Discussion

### 6.4.1 Comparison with ABR

The comparison between ABR wave V and ASSR is illustrated in figure 6.5. Both curves present a similar trend of decreasing latency with increasing frequency. A two-way ANOVA test was run to compare the data from ABR and ASSR, the test examines the effect of independent factors on the latency estimate. The independent factors are the frequency ( $n = 6$ ) and the measurement technique ( $n = 2$ , ABR or ASSR). This test compares the slope of each curve. Results are declared significant if the  $p$ -value is less than 0.05. It results that  $p = 0.079$ , there is therefore no significant difference between ABR and ASSR latency estimates, this confirms that their slope are close to

each other. The difference between ABR and ASSR is thus only a vertical shift of about 10 ms.

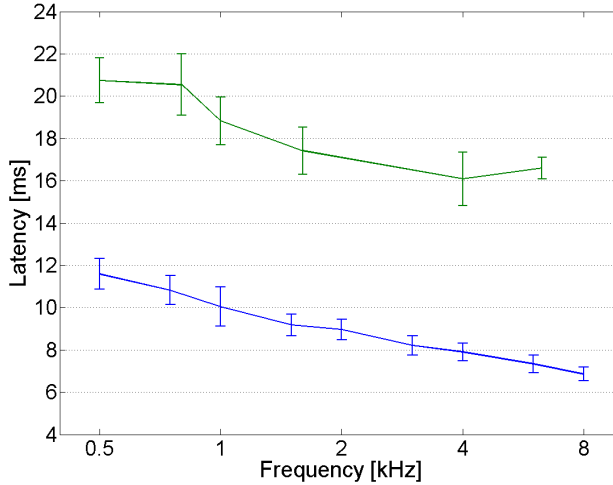


Figure 6.5: Comparison between latency estimates obtained from ABR recordings (wave V, blue curve) and ASSRs (green). The ASSRs are about 10 ms longer than ABRs and their inter-subject variability is also higher.

Compared to ABRs, ASSRs have a higher intra-subject variability, the maximum variation for ABRs is 1.13 ms against 1.9 for ASSRs. The wave V could be detected in most of the subjects in the three different runs whereas the ASSRs appeared harder to detect due to the high noise level, this intra-subject variability is based on only two runs and in fewer subjects. It should be noted that at least 3000 epochs were recorded for ABR and only 600 for ASSR, this might also explain why ABRs appeared less noisy than ASSRs. It seems therefore that, in the present investigation, ABRs are more reliable than ASSRs, i. e. less susceptible to noise, especially at higher frequencies.

Their inter-subject variability also differs and it is much higher for ASSRs than for ABRs. It should be reminded that the groups of subjects used for both experiments are not the same and it seems that the group used for ASSRs presents a larger variability. Another reason could be the origin of ASSRs. Contrary to the wave V which is generated in the brainstem, ASSRs are the sum of sources located at different

levels of the auditory pathway, that is brainstem and higher stages such as the auditory cortex. At this higher level, the processing is greatly affected by the state of arousal (see Chap. 2) which differs from one subject to the other. This can explain why ASSR has a greater variability. The most striking difference between ABR and ASSR latency estimates is their actual values, ASSRs are about 10 ms longer than ABR. The modulation frequencies used (83, 88, 93 Hz) ensure, according to the literature, that ASSR is mainly generated in the brainstem (albeit high brainstem), so that the latency should be of the same order as ABRs. Both ABR and ASSR data are each in good agreement with previous studies, so any experiment-related error can be discarded. Other investigators have also noticed this latency difference between ASSRs and ABRs (John and Picton, 2000; Rance et al., 2006) and a few hypotheses have been suggested (Purcell et al., 2006). It could be that the main contributor to ASSR is located above the wave V generator, that is above the inferior colliculus. Purcell et al. (2003) derived the ASSR delay above the BM (corresponding to  $\tau_{synaptic} + \tau_{neural}$  in Fig. 4.3) and found a value of 15.3 ms at 3.6 kHz. The value obtained in the present ABR experiment for the delay above the BM ( $\tau_{synaptic} + \tau_{neural}$ ) is 4.5 ms (calculated from table 4.1). There is a difference of 10.8 ms, which would confirm that ASSR originates dominantly from areas above the ABR generator. As Purcell et al. (2003) concluded, "cortical and/or thalamus sources may play a significant role (...) in addition to sources in the colliculus and lemniscus" for the 85 Hz ASSRs. Previous research have tried to localize the sources of 80-Hz range ASSRs, Herdman et al. (2002) used EEG and found the largest activity in the brainstem whereas Schoonhoven et al. (2003) used magnetoencephalography (MEG) and located the sources in the primary auditory cortex. The precise source of the 80-Hz ASSR remains unknown but the possibility a dominant upper-brainstem generator (inferior colliculus and beyond) is not excluded. This could explain the delay between ABR and ASSR latencies. Furthermore, higher stages of the auditory pathway are more affected by the state of arousal of the subject, since this state changes from one subject to the other, this could explain why the inter-subject variability is larger for ASSRs than for ABRs.

It should also be noted that the exact source for the absolute ASSR delay is not well understood, and the main interest is rather the latency difference between high and low frequencies. For ASSRs, the difference between 0.5 and 6.3 kHz is 4.1 ms and

4.7 ms for ABRs between 0.5 and 6 kHz. These two values are very close and confirm the fact that ABRs and ASSRs reflect the tonotopic organization of the cochlea in the same way. The latency difference between ABRs and ASSRs appears therefore mainly at a neuronal stage and not at a cochlear level, meaning that ABRs and ASSRs might not be generated at the same place. The ASSR could correspond to the Na/Pa complex of the MLR, which occurs around 20 ms, as illustrated in Fig. 2.3.

A part of the latency difference between ABRs and ASSRs could possibly occur at a cochlear level, the notion of filter build up time was introduced in Fig. 2.5, it corresponds to the time needed to go through the active process of the cochlea. Neurons start firing at different time during the filter build-up time  $\tau_{filter}$  (John and Picton, 2000) and Eggermont (1979b) introduced a factor  $\beta$  to account for this, with  $\beta = 0.5$  for transient stimuli and  $\beta = 1$  for AM-tones (John and Picton, 2000). This suggests that ASSRs occur at the peak of the BM filter (requiring the entire filter build up time) whereas neuron synchrony leading to TBABR occurs before this peak. This difference is hard to quantify but may explain part of the latency difference observed between ASSRs and ABRs.

One of the striking result from the comparison between OAEs and ABRs in Chap. 5 was the breakdown at 2 kHz. It appeared that below 2 kHz the latency of OAEs and ABRs as a function of frequency have similar slopes and that above this limit the curves diverge. As mentioned earlier, the ANOVA test showed that the ABR and ASSR curves have similar slopes, this means that the ASSR data would also present some divergence with the OAE data above 2 kHz. This is only a rough approximation since the subjects are different, a comparison on an individual basis could bring a more precise idea of the cochlear mechanics.

### 6.4.2 Gender difference

An interesting observation concerns subject 4 whose latencies are lower than the other subjects. The experiment was repeated on this subject and it confirmed the values previously found. The possibility of measurement error can therefore be discarded. These lower latencies can be partially explained by the gender of the subject, who is a female. As Fig. 6.6 shows, ASSR latencies are shorter for females, as it was also measured by John and Picton (2000). Fig. 6.6 illustrates an almost constant

difference of about 1 ms between men and women. To gain a better understanding of the behaviour of subject 4, further tests (e.g. ABR) could be carried out to find out if she also has lower wave V latencies. This would indicate whether her neural processing is faster than for other subjects.

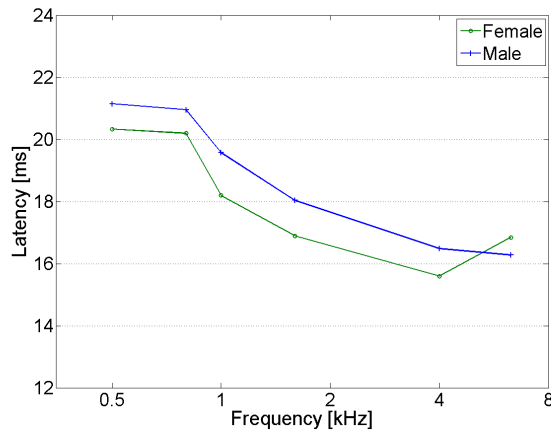


Figure 6.6: ASSR latency as a function of frequency for female and male subjects.

## 6.5 Conclusion

Until now, no published research had compared ASSR and ABR latency in normal-hearing adults. The comparison from the present study concludes that: a) ABRs are a more reliable recording technique than ASSRs (better intra-subject reliability), b) the noise level is lower at high frequencies for ABR, c) ABRs and ASSRs present the same decrease of latency as a function of frequency ( $p > 0.05$ ), d) ASSRs in the 80-Hz range are probably generated at a higher stage than ABRs, i.e. inferior colliculus and above. This difference in generation locus explains the latency difference between ABRs and ASSRs. This may explain, to some extent, the higher inter-subject variability observed for ASSRs, which are more affected by arousal state. An interesting topic for future research would be to compare ABRs and ASSRs in

individuals as this was the case in Chap. 5 between ABRs and OAEs. This could bring a better understanding of the different delays involved in ABR and ASSR generation and provide a better knowledge of the origin of ASSRs. Controlling the state of arousal of the subjects would also decrease the effect of higher neural stage and could provide a more reliable comparison between ASSRs and ABRs. Future investigations could also estimate cochlear delays,  $\hat{\tau}_{BM}^{(ABR)}$ , from ASSRs and compare it directly with the OAE delay.

# Chirp-evoked otoacoustic emissions

---

### Abstract

Chirps are stimuli used to compensate for the cochlear dispersion and thus theoretically create a simultaneous maximum displacement of the entire BM. They have been traditionally used to evoke optimised ABR or ASSR recordings, making the responses more detectable than when using clicks. This obviously has clinical advantages. The present experiment explains how a chirp was used to compensate for the OAE travel time and the effects of such a chirp on OAEs are investigated. Clicks and chirps were presented at 40 dB peSPL and an analysis of the recorded responses was made in the time, frequency and time-frequency domains. Results of two exemplary subjects are used to demonstrate that it is possible to elicit OAEs from a broad range of frequencies and synchronise their arrival time in the ear canal (EC). The responses obtained are concentrated in time and sit well above the noise floor. The present experiment represents an important first step in the study of chirp-evoked OAE and further research could verify the usefulness of such a method for clinical purposes.

### 7.1 Introduction

The previous chapters presented different physiological measurements of BM latency. OAEs and ABRs were measured using tone bursts as stimuli and ASSRs were evoked by AM tones. Each of these methods could be used to estimate a latency-frequency function for each subject tested. The trend of this function is similar across methods, i. e. the latency decreases with increasing frequency, reflecting the tonotopic

organization of the cochlea. In Chap. 5, it was shown that this function can be expressed as a power law function of the type  $t = bf^{-a}$  with  $a < 1$  and  $b > 0$ . Each of the physiological measures used represent an indirect measurement of latency, with its source located at different stages in the auditory pathway. ABRs are responses generated at a brainstem level, 80-Hz ASSRs are generated at a high brainstem level and only OAEs are generated on the actual cochlear level.

The time difference between the tone-bursts OAEs (TBOAEs) generated at low and high frequencies is estimated at around 12 ms between 0.5 and 8 kHz. It is therefore, in principle, possible to create a stimulus that takes this difference into account. When evoked by such a stimulus, the OAE at 0.5 kHz should be generated 12 ms before the OAE at 8 kHz, so that they reach the stapes at the same time. The energy hence recorded in the EC should be the sum of all the OAE components, therefore reaching a higher response amplitude than if OAEs were delayed.

Such a stimulus is called a chirp and historically, chirps have been used to compensate for the cochlear delay and improve the detection of physiological signals such as ABR, ASSR and post-auricular-muscle response (PAMR). The chirps used were based on different estimations of the cochlear delay: for Dau et al. (2000), the change in the travelling wave velocity along the cochlear partition was obtained from the linear cochlea model of de Boer (1980), which was combined with the frequency-place mapping of Greenwood (1990). This allowed to compensate for the difference of maximum displacement between high and low-frequency regions of the cochlea and lead to an increase of the ABR magnitude. The chirp presented in Dau et al. (2000) was subsequently used in other ABR studies (Bell et al., 2002; Wegner and Dau, 2002; Junius and Dau, 2005) or PAMR recordings (Agung et al., 2005). Fobel and Dau (2004) compared chirps based on SFOAE and ABR data with the de Boer model-based chirp. They showed that ABR recordings can be optimised, i. e. obtained with a higher response in a shorter recording time, when a chirp based on ABR data is used. The idea behind such a chirp is to excite the BM in a way that the auditory nerves fire simultaneously, leading to a higher neural synchrony and giving higher ABR amplitude. Elberling et al. (2007) also found that ASSR recordings can be optimised by using chirps based on derived-band ABR and narrow-band compound action potentials from Don et al. (2005) and Eggermont (1979a), respectively.



The reason why so many studies have focused on improving AEPs recordings is certainly their widespread clinical use. Finding the best stimulus that can improve audiological diagnostics is of great importance and since OAE recordings are part of the clinical tests routinely used to assess patients' hearing, it seems necessary to investigate the possibility to optimise them, too. The optimal stimulus to record broadband ABR is a chirp based on ABR data, so the optimal stimulus to record OAEs could be a chirp based on OAE data. Only the study of Fobel and Dau (2004) actually investigated the use of a chirp based on OAE data (O-chirp), but this was used to record ABRs.

The O-chirp was generated to excite the whole BM simultaneously and it is based on the assumption that  $\tau_{BM} = \frac{\tau_{OAE}}{2}$ , which was shown to be incorrect at high frequencies (Chap. 5). Furthermore, what is needed in the present study, is to excite the BM so that the OAEs reach the stapes at the same time, i. e. the backward travelling time should be taken into account. The stimulus commonly used for clinical test is a click and the difference between a click evoked OAE (Click-OAE) and a chirp-evoked OAE (Chirp-OAE) is schematically illustrated in Fig. 7.1.

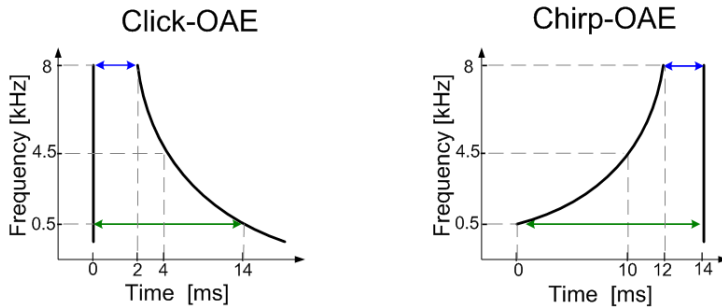


Figure 7.1: Expected OAE responses when the stimulus used is a click (left) and an OAE-based chirp (right). These schematic plots use a time-frequency representation of the signals. In this example, the 8 kHz OAE is recorded 2 ms after the stimulus (blue arrows) and the 0.5 kHz OAE has a latency of 14 ms (green arrows).

Fig. 7.1 shows the time-frequency representation of ideal recordings. At time  $t = 0$ , the stimulus appears (click or chirp) and the actual OAEs appear later. For the Click-OAE (left), the stimulus contains a broad frequency range concentrated in a short time slot, the high frequency regions are excited first and the OAEs from these regions are

therefore recorded first. In this example, the 8 kHz OAE is recorded 2 ms after the stimulus (blue arrows) and the 0.5 kHz OAE has a latency of 14 ms (green arrows). This latency difference is represented in the time-frequency domain by a downward chirp. If a chirp is used as a stimulus, Fig. 7.1 (right), the time between stimulus and response at a certain frequency stays identical to the click case. The stimulus can be seen as the "mirrored" version of the Click-OAE curve on the left. The expected curve for the OAE responses in the case of the chirp is a straight line.

If the predictions for the Chirp-OAE are correct, this type of stimulus would be of great interest for clinical use. The response hence elicited is somehow concentrated in time, which increases the amount of energy available in the ear canal and makes, therefore, the detection of the OAE signal easier.

The following section presents in more detail how the OAE-based chirp is generated and describes the experiment carried out to record Chirp-OAEs.

## 7.2 Methods

### 7.2.1 Subjects and procedure

Transient-evoked OAEs were recorded from seven normally hearing subjects. An RME Fireface 800 24 bit A/D-D/A converter was used for data generation and acquisition, together with a Tucker-Davis PA5 programmable attenuator to control the levels. The stimuli were generated and played in MATLAB using the open source pawayplay set of dlls written by Matt Frear at the University of Western Sydney. These dlls were created with the use of the PortAudio Portable Real-Time Audio Library available under the Open Source Initiative MIT license. The stimuli were presented to the test subject via an Etymotic research ER-2 probe and recorded via an ER10B+ precision microphone with 40 dB gain, and then bandpass filtered (0.6-5 kHz) via an analogue Rockland 852 filter. The noisiest 10% of all recorded epochs were removed by an artifact rejection template and the remainder averaged. Test subjects were screened to ensure all had pure-tone thresholds better than 15 dB HL in the range 0.25-8 kHz.

Two sets of stimuli were presented to each subject, firstly clicks of approx.  $113 \mu\text{s}$  duration and secondly upwardly swept chirps designed to compensate for OAE travel time. All stimuli were presented at 40 dB peSPL calibrated in an IEC711 coupler. Two times 4000 stimuli of each type were presented to the left ear of each subject. Each stimulus epoch lasted a total of 99 ms.

### 7.2.2 Stimuli

The dispersion in the cochlea will separate frequency components in time in the recorded TEOAEs when simple rectangular transient stimuli are used. In this study, the dispersion in OAE components is investigated by compensating fully for this dispersion. OAE delay or latency as a function of frequency has been measured in Chap. 5 in eleven normal hearing adult subjects. Based on the delay estimates obtained from this study, an empirical relation was established (inspired by Gorga et al. (1988) for ABR delay estimates):

$$\begin{aligned}\tau(f) &= 10^{a \log_{10}(f)+b} \\ &= 10^b f^a\end{aligned}\tag{7.1}$$

where  $\tau(f)$  is the latency as a function of frequency  $f$ ,  $a$  and  $b$  are parameters fitted to the experimental data in Chap. 5,  $-0.574$  and  $0.8254$ , respectively. This gives an empirical relation between delay and frequency for OAEs, averaged across subjects, at the level of 66 dB peSPL. Making a change of variables  $\tau \rightarrow t_0 - t$  to invert the time relation, the instantaneous frequency  $f(\tau)$  needs to be solved for from this empirical formula in order to derive a chirp function capable of compensating for cochlear delay. It is possible to show that the instantaneous frequency is given by:

$$f(t) = 10^{-\frac{b}{a}} (t_0 - t)^{\frac{1}{a}}\tag{7.2}$$

Integrating over time yields the instantaneous phase,  $\phi(t) = 2\pi \int_0^t f(\tau) d\tau$ , given by

$$\phi(t) = -2\pi 10^{-\frac{b}{a}} \frac{a}{a+1} \left( (t_0 - t)^{\frac{a+1}{a}} - t_0^{\frac{a+1}{a}} \right)\tag{7.3}$$

where the constant of integration was chosen to ensure zero phase at time  $t = 0$ . Finally, the chirp function is obtained from:

$$x(t) = A \sin \left( -2\pi 10^{-\frac{b}{a}} \frac{a}{a+1} \left( (t_0 - t)^{\frac{a+1}{a}} - t_0^{\frac{a+1}{a}} \right) \right) \quad (7.4)$$

Similar to Dau et al. (2000), it was desired that the chirp stimulus has a flat magnitude spectrum to maximise comparison with the rectangular stimuli. Since the instantaneous frequency changes slowly at low frequencies relative to high frequencies, the power spectra would be dominated by the low frequencies. This has been roughly compensated for by varying the amplitude as a function of time, i.e.  $A \rightarrow A(t)$ . It can be shown that allowing the amplitude to vary as:

$$A(t) \propto (t_0 - t)^{\frac{1-a}{2a}} \quad (7.5)$$

will result in a flat magnitude spectrum.

The chirp used in the present experiment is designed to have a flat spectrum between 0.5 and 4.5 kHz. The reason for limiting the chirp to 4.5 kHz is related to the difficulty of detecting high frequencies OAE. As experienced in the TBOAE experiment (Chap. 3), high frequency TBOAEs have very short latencies, often merged with the stimulus. It is therefore difficult to extract the signal from the recorded signal. Limiting the chirp to 4.5 kHz increases the chance to unambiguously detect the response. The chirp also has a 0.5 ms lead in/out hanning window applied to help minimise startup and offset transients. The normalised time series for the empirically determined chirp running from 0.5 to 4.5 kHz is shown in the bottom trace of figure 7.2(a). The trace above this shows the click (top). Figure 7.2(b) shows the power spectra for the two stimuli types. It can be seen that, within the bandwidth of interest, the magnitude spectra are flat and comparable across stimuli types.

### 7.2.3 Time-frequency distribution

Otoacoustic emissions, evoked from the transient (click and chirp) stimuli used in this study, are inherently non-stationary signals. Essentially their spectral content varies with time, due to the dispersion in the cochlea. In order to analyse such signals, non-

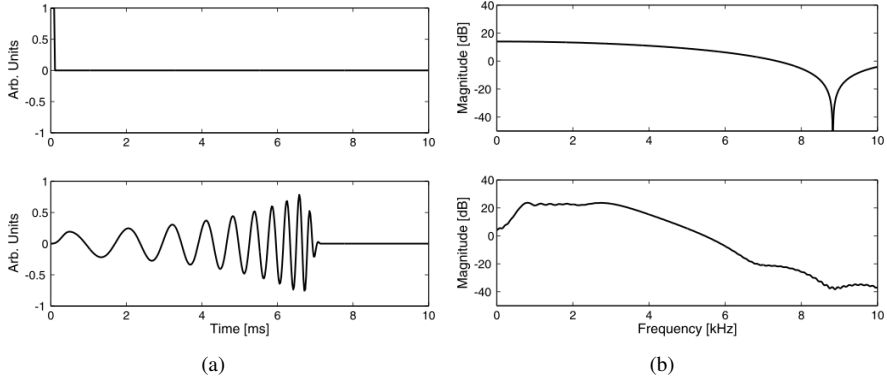


Figure 7.2: Stimulus time series (a) and spectra (b): top - transient, bottom - chirp. As expected the transient signal has a very broad spectrum and the chirp was generated to stay flat on a limited region, from 0.5 to 4.5 kHz.

stationary methods must be applied, namely a time-frequency representation. There are a number of examples in the literature for using time-frequency methods with OAEs (Probst et al., 1986; Tognola et al., 1997; Sisto and Moleti, 2002; Konrad-Martin and Keefe, 2003; Jedrzejczak et al., 2005; Moleti et al., 2005). These examples have typically used complex mathematical techniques. In this preliminary study, simplicity is sought, therefore, only the spectrogram will be used. The spectrogram is defined as the squared modulus of the short-time Fourier transform, given here in discrete form:

$$S_x(n, f) = \left| \sum_{m=-\infty}^{\infty} x(m)w(m-n)e^{-2\pi i f m} \right|^2 \quad (7.6)$$

where  $x(m)$  is the signal and  $w(t)$  is a windowing function. The window used in this study was chosen to be a Hanning window,

$$w(n) = \frac{1}{2} \left( 1 - \cos \left( \frac{2\pi n}{N-1} \right) \right), \quad 0 \leq n \leq N \quad (7.7)$$

where  $N$  is the length of the window in samples. Here this was chosen to be 201 samples long, yielding a duration of approximately 4.19 ms. This was chosen as a

compromise between spectral and temporal resolution. At 4 kHz, OAE delay is of the order of 4 ms. Therefore if the window was longer, yielding a better spectral resolution, then the onset of the OAE at this upper frequency would be impossible to detect. The energy of the stimulus which is approximately 40 dB greater than the OAE would overwhelm the results.

Returning to the spectrogram, multiplication in the time domain by a windowing function yields a convolution in the frequency domain, therefore the properties of the window used are very important. It can be shown that the Hanning window chosen has a 3 dB bandwidth of approximately 330 Hz, which is therefore the basic limit to separate closely spaced frequency components. The highest side lobe ( $-32$  dB below the main lobe) should be sufficiently small that it does not corrupt or further limit the frequency resolution, hence the use of the Hanning window.

## 7.3 Results

Among the seven subjects that participated in this experiment, only four had recordings suitable for analysis. The other three subjects showed a high noise level, with an OAE signal difficult to detect and their overall results were ambiguous. The present section details the results obtained in two exemplary subjects, who showed very good OAE signals.

### 7.3.1 OAE time series

#### Short latency

Figure 7.3 shows the short latency (top panels), defined here as 0-10 ms after stimulus onset, and the long latency (bottom panels), 5-40 ms after stimulus onset, for subject scf for the click (a) and the chirp (b). The figures are all plotted in pascals along the ordinate and milliseconds along the abscissa. In general, the results from the two sub-averages (3600 epochs in each) for each stimulus type show excellent agreement for this test subject.

The short latency results for the click show the 1 ms lead in time, due to the ER-2 earphone delivery tube. The maximum peak-to-peak pressure in the response is of the

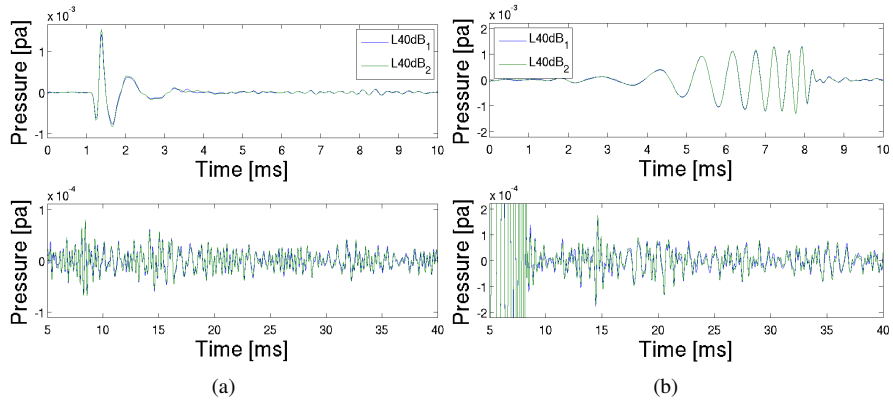


Figure 7.3: Subject scf: (a) click responses and (b) chirp responses, at 40 dB peSPL. Two sub-averages are plotted in each case (blue curve and green curve). The superimposition of these two curves proves the very good repeatability of the data. The top plots are short latency and the lower plots long latency responses.

order of 2.25 mPa. The chirp response has a maximum peak-to-peak pressure of the order of 2.8 mPa. Since these two values are close to each other, the calibration in terms of peSPL was carried out correctly. Figure 7.4 shows the short latency power spectra for the same subject, scf. Across stimuli types there is an overall level disparity in the 1-4 kHz region of interest for this study. The click response is approximately 10 dB lower in magnitude than the chirp. Therefore our calibration of the stimuli in terms of peSPL may not be the best method for these stimuli types. This makes the OAE results hard to compare in terms of magnitude. The chirp stimulus is a more powerful evoker of a response than the click used in this study. Future work will be carried out to ensure that the stimuli are calibrated correctly so that the sound pressure levels in the ear canal are equal. However, the chirp stimulus has the correct flat spectrum between 1-4 kHz, making the comparison possible in terms of timing and waveform shape, if not directly in magnitude.

Looking at figure 7.4(a), there exists some ripple in the 3-4 kHz region of the click response short latency spectrum. This is due to the definition of short latency taken here between 0 and 10 ms. Inspecting the lower plot in figure 7.3(a), which shows the OAE, there is clearly OAE energy in the region 5-10 ms. Thus, the magnitude spectra plotted here, reflect the addition of the stimulus and the OAE components in

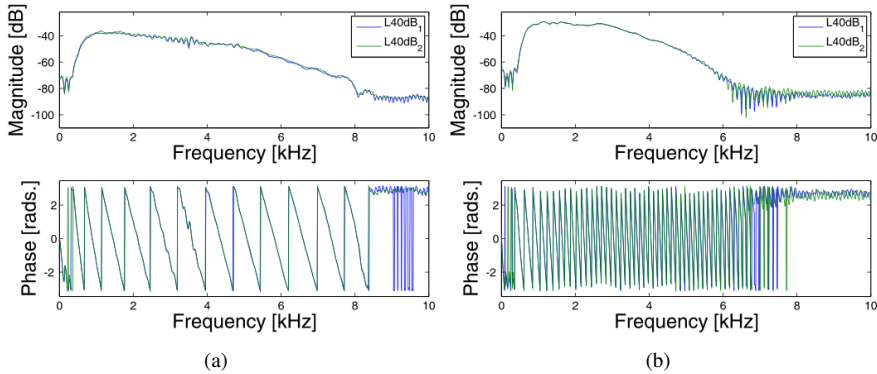


Figure 7.4: Subject scf, power spectra for (a) click responses and (b) chirp responses, at 40 dB peSPL.

this frequency range, their respective phases will result in the ripple seen on the power spectra. This is not observed in the chirp response spectrum as the OAE component does not dominate the time series until later (around 12-16 ms), due to the phase delay introduced by the stimulus itself.

### Long latency

For subject scf, all of the stimuli types resulted in strong evoked-OAEs and synchronised SOAEs as well. These are shown in the lower plots of figure 7.3, in the range 5-40 ms. This subject demonstrated strong SOAEs throughout the entire analysis window (approx. 100 ms). The click results, in Fig. 7.3(a), show a broadband component sweeping down in the 5-15 ms time frame as one would expect from a transient OAE. It is possible, throughout the time series, to see regions of high and low frequency. This is not entirely common to observe in all test subjects, as this subject had particularly powerful OAEs. Regions of high-frequency bursts occur at 7-10, 12-13, 20-23 and 25-28 ms in this case. Low-frequency regions at 10-12, 16-18, 23-24 ms. The location of the first and most powerful burst is at 7-9 ms. This could correspond to internal reflection and multiple re-emission from the cochlea

The chirp OAE, in figure 7.3(b), shows a clear transient like response around 14-15 ms. This is exactly what one might expect if the OAE latency were correctly



compensated for, and multiple or a broad range of frequencies would arrive at the recording microphone at the same time. After this, the OAE pattern becomes similar to the click response.

Subject jjm shows similar long latency results, as shown in figures 7.5. For this subject as well, the repeatability is excellent across runs. For this subject, similar patterns are seen in the OAE data across stimuli types, where the chirp shows an apparent narrower time range for its first burst of energy.

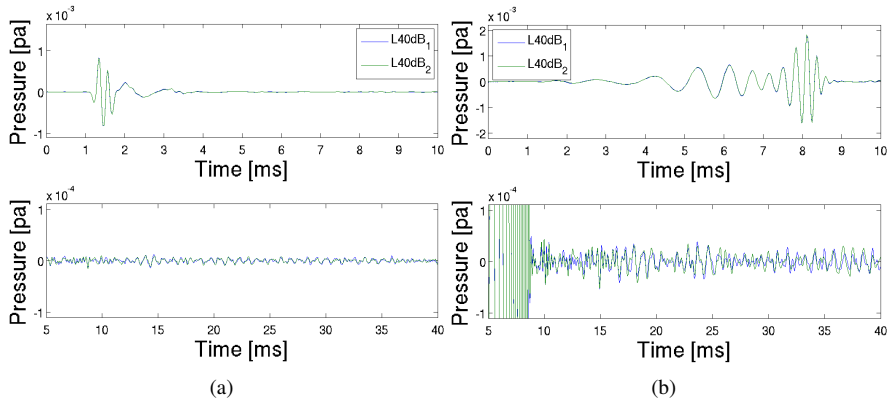


Figure 7.5: Subject jjm: (a) click responses and (b) Chirp responses, at 40 dB peSPL.

### 7.3.2 Time-frequency analysis

An analysis in time and frequency was performed for each subject, according to section 7.2.3. The results for subject scf are shown in Fig. 7.6. The shape of the OAE differs between the click and the chirp and their time-frequency representations recall the ones schematically presented in Fig. 7.1. For the click response, Fig. 7.6(a), the stimulus clearly dominates the distribution in the 0-5 ms region. There is a broad spectrum in this time frame between 0.5-6 kHz (upper limit of the figures). A small amount of dispersion can be seen at low frequencies, where there is a small delay relative to the high frequencies. This is due to the transducers used, the subjects ear canal and middle ear, and the analogue filtering before data acquisition. The first OAE response can be seen after the stimulus, as a downward chirp of energy, with a

maximum level about 35 dB below the highest stimulus level. Between 1.2 and 4 kHz it is possible to trace a line within the temporal window of 8-18 ms. The magnitude is not constant throughout due to the random nature of the local reflectors on the BM.

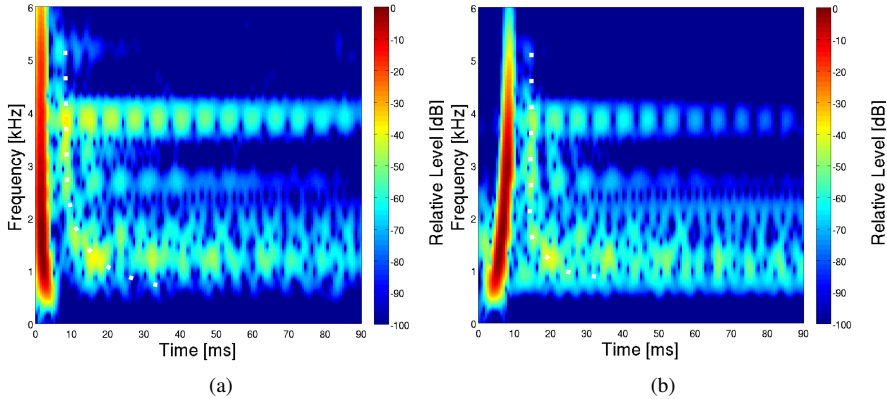


Figure 7.6: Subject scf: Spectrogram for (a) click responses and (b) chirp responses, at 40 dB peSPL. The stimulus can be seen in the left part of each spectrogram, recognizable by its high level. OAEs occur later, about 40 dB below the stimulus level. The white dots indicate the trend of the recorded OAEs.

For the chirp, figure 7.6(b), the dispersion of the chirp can be seen in the 1-10 ms region. The stimulus has a broad spectrum from 1 to around 5 kHz, which is narrower than the click spectrum and this corresponds well to the way the chirp was generated (see section 7.2). The shape of the response also differs from the click-OAE. For the chirp, the OAE responses are more aligned, forming almost a straight line at 15 ms, between 1.6 and 4 kHz. Between the click and the chirp, the energy of the response is shifted in time. Only the low frequency OAEs (1-1.6 kHz) are not perfectly aligned with the other responses, for the reasons mentioned earlier for the click (transducer, EC and middle ear). For a given frequency, the delay between stimulus and response stays identical in both the click and the chirp, as expected under the coherent reflection filtering theory. The effect of the chirp is to synchronise the OAE, therefore resulting in a click-like response, as observed in the time series Fig. 7.3(b) around 14-15 ms.

At 4 kHz, in figure 7.6, a band of energy beats with a periodicity of around 4-5 ms. Two hypotheses can explain this phenomenon:

- Internal reflection and re-emission of an OAE component, as the total travel time at 4 kHz is approximately 4 ms on average. There also appears to be a reduction in magnitude of the component as time goes by, consistent with energy dissipating in the reflection and re-emission process.
- Two or more dominant SOAE components exist closer in frequency than the spectral resolution of the spectrogram. Two closely spaced sinusoids add up to form a beating wave, with the periodicity of the beating being proportional to the frequency difference.

Further analysis would be needed to differentiate between these two hypotheses.

The results for subjects jjm are shown in Fig. 7.7. Similar observations can be made since the stimuli and response patterns are close to the ones observed for subject scf. However, the response magnitudes are much reduced. This subject represents a more typical subject, while scf had particularly strong OAEs. The chirp response of subject jjm presents a greater alignment of the first OAE reflection, compared to subject scf.

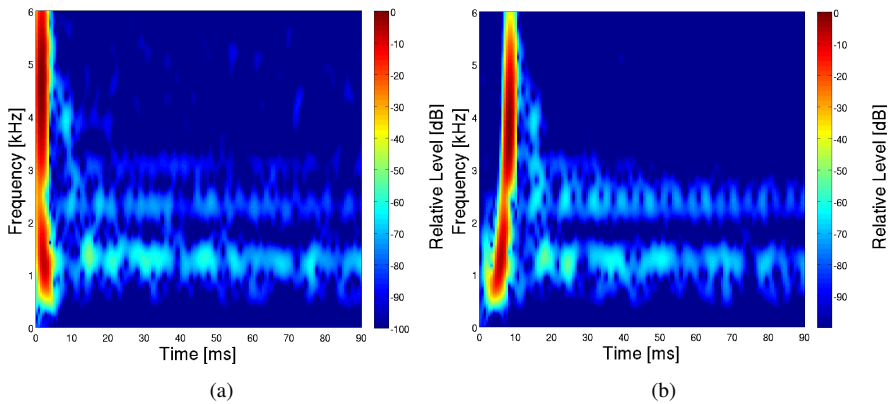


Figure 7.7: Subject jjm: Spectrogram for (a) click responses and (b) Chirp responses, at 40 dB peSPL.

## 7.4 Discussion and conclusion

The goal of this experiment was to investigate the possibility to use an OAE-based chirp as a stimulus to elicit OAEs. The compensation of the OAE travel time was based on the TBOAEs that were recorded in a previous experiment. The results of the present experiment showed that the chirp could compensate for OAE delay and helped to synchronise the OAE responses. When analysed in time and frequency, the pattern of the chirp evoked-OAE was close to a straight line between 1.6 and 4.5 kHz.

The experiment presented here was only intended as a preliminary investigation of the chirp-evoked OAE. The promising results imply that there is potential in such paradigm. There is also space for improvement, such as compensating for the inherent greater dispersion at low frequency by modifying the chirp parameters  $a$  and  $b$  in Eq. 7.1. Since the chirp was based on OAE recordings made at 66 dB peSPL, a chirp based on lower level OAE recordings could also be used. This study also showed that the chirp had a higher spectrum magnitude than the click, hence making the comparison slightly biased. An improvement would be to calibrate the stimuli using their spectrum and not only time information.

The present experiment showed that, in a similar way as ABR responses were maximised, OAEs can also be potentially optimised using a chirp. The response obtained by using a chirp based on OAE data is narrower in time compared to a click response. In this way, the response sits well above the noise floor, which helps for its detection. Further work is, however, needed to find the optimal stimulus to evoke OAEs and clinical tests could be performed to confirm the usefulness of such a chirp.

# Overall discussion

---

This final chapter presents the overall discussion for this thesis. In the first section, the main experimental results are summarized. This is followed by a discussion of the results and a comparison of the three measurement methods used, OAE, ABR and ASSR. As a conclusion about these techniques, advice is given regarding the best way to derive cochlear delay from non-invasive measurements. The following section then outlines the implications of the present findings for cochlear modelling and suggests projects that could be conducted in the future.

## 8.1 Summary of the main results

The main contribution of this work is the comparison of three techniques to measure auditory delays in human. This comparison was done for the same subjects when measuring OAE and ABR and on different subject groups for ABR and ASSR. Chap. 3 presented how to estimate TBOAE latency with the help of a new paradigm to detect the OAE burst. This successfully lead to individual estimates of TBOAE latency that were in agreement with previously published data. This experiment also studied the effect of noise on OAE latency and demonstrated that the effect is negligible. However, when the SNR is fixed at a given frequency, the number of averages needed is greatly decreased. There is thus an important gain in time when using such a criterion, compared to having a maximum number of averages. The third important result from this experiment is the use of individual data instead of averaged data. It was shown that the intra-subject variability was small compared to the inherent inter-

subject variability. This is of particular interest when applying this measurement in a clinical environment.

The ABR experiment in Chap. 4, which estimated TBABR delays, also proved to be very reliable in terms of low intra-subject variability and was in good agreement with previous studies. The difference in the nature of the stimuli (rise time and level) accounted for the differences from other studies. The ABR experiment also showed that the inter-subject variability is frequency dependent, being lower at high frequencies. It was concluded that this was due to a better detectability of the wave V at these high frequencies. The delay between wave III and wave V was also calculated; it confirmed that this delay is independent of frequency, but some variation across subjects appears. This interpeak delay was used to estimate the BM delay from the ABR latency, denoted  $\hat{\tau}_{BM}^{(ABR)}$ .

Chapter 5 presented the comparison between  $\hat{\tau}_{BM}^{(ABR)}$  and the OAE delay  $\tau_{OAE}$ . Such a comparison had only been investigated before on a limited frequency range. Here, the study was carried out between 0.5 and 8 kHz. This is a crucial point since the properties of the BM change with frequency. The results obtained in the current study support the coherent reflection filtering theory for OAE generation for frequencies below about 2 kHz; at these frequencies, the OAE delay is believed to be twice the BM delay ( $\tau_{OAE} = 2\hat{\tau}_{BM}^{(ABR)}$ ). This confirms that the OAEs use the same path for their inward and backward travel in the apical part of the cochlea. For the more basal part, the experiment showed that a round-trip does not seem to exist anymore. Indeed, for frequencies above 4 kHz, the relation becomes  $\tau_{OAE} = \tau_{BM}$ . The present study hypothesises that for high frequencies, OAEs travel back through both the fluid and the cochlear partition. The fluid propagation being faster, the backward travelling wave reaches the stapes earlier and stronger. This mode of propagation is dominant for high frequencies where the cochlear amplifier is more active. On the contrary, for low frequencies, this fluid propagation is hindered by irregularities on the BM or even cancelled by signals generated close to the CF place. The dominant propagation at low frequencies is therefore through the cochlear partition.

A third technique was used to measure delays in the auditory system. Chap. 6 presented how ASSR latency can be estimated. Until now, no studies had compared the ABR and ASSR latency in normal-hearing adults. In the present study, ASSR in

the 80-Hz range had delays about 10 ms above the average ABR delays. This casts doubt on the brainstem origin of the ASSR. Contrary to the theory, it seems that ASSRs evoked by 80-Hz modulated tones are generated by sources located above the inferior colliculus. The latency difference between high and low frequencies is similar for ABRs and ASSRs, around 4.5 ms between 0.5 and 6 kHz. These two measurements therefore reflect the cochlear processing, but at different stages of the auditory pathway.

Finally, an experiment was conducted to investigate OAEs evoked by a chirp. This chirp was generated using the OAE data obtained in Chap. 3. Only the data of few subjects could be analysed and the results showed that the chirp-evoked OAEs were synchronised in time with a higher amplitude than click-evoked OAEs. This experiment demonstrated that the OAE-based chirp developed here is able to compensate for the OAE travelling time and that the responses reach the ear canal simultaneously.

## 8.2 Discussion

The combination of the three experimental methods brings a unique view on auditory delays in humans. OAEs, ABRs and ASSRs can be recorded with different techniques, each having advantages and drawbacks. A detailed comparison is made in the following:

**Reliability:** This is measured by re-testing on the same subject and is quantified by the intra-subject variability. As shown in the previous chapters, OAE measurements have the lowest variability, followed by ABR and ASSR. OAE signals are the most constant signals in time.

**Robustness to noise:** This study also showed that OAEs are not as sensitive to noise as ABR and ASSR can be. The noise recorded for OAE comes from the room where the experiment takes place and also from the subject (body movements and heartbeats). For evoked potentials, these noise sources add to the neurophysiological noise. It is easier for a subject to stop moving than to stop thinking, making the background physiological noise higher for AEPs.

**Recording duration:** OAE measurements were the fastest with eight tone-burst frequencies recorded in about 45 min. ASSRs took about one hour for six frequencies but need a longer preparation time. The ABR recording required about two hours for eight TB frequencies; this was necessary to increase the SNR and obtain more detectable signals. Besides, the OAE experiment can be shortened if a fixed SNR level is used at each TB frequency. As explained in chapter 3, the OAE signal can be detected with just a few hundred averages.

**Experimental setup:** The OAE recording was the easiest to set up. Only a microphone and a speaker in a sealed ear canal are needed. The preparation is longer when recording evoked potentials, where electrodes need to be positioned on the subject's head. Although ABRs and ASSRs can be recorded using the same setup, this was not the case here due to the different locations of the laboratories where the experiments were conducted. ASSR did not require the use of an electrode cap which made the preparation shorter than for the ABR recording. There were more parameters to change and adjust prior to the ABR experiment than before the ASSR experiment.

**Compatibility with previous studies:** The previous chapters have shown that all three measurement techniques gave results in good agreement with published studies.

From this comparison between OAEs, ABRs and ASSRs, it appears that OAE measurements are the most effective, combining reliability, rapidity and ease of setup. However, the choice of the measurement method depends on the purpose of the study. One should keep in mind that cochlear delays cannot be estimated directly from ASSR recordings due to the uncertainty about their generation sites. ASSR can, however, be used to confirm delays estimated from ABR data. A great advantage of ABR is the limited involvement of the cochlear amplifier in their generation. Contrary to OAE, the active mechanism is not dominant in ABR generation. The delays, hence estimated, reflect more the passive behaviour of the BM. Besides, the subtractions made to estimate cochlear delay from wave-V latency are rather simple. Although OAE measurements seem the most reliable, they might not represent the best choice if one wants to measure cochlear delays. The OAE generation mechanisms are still not fully understood. Only for certain frequency ranges, the relation between  $\tau_{BM}$  and



$\tau_{OAE}$  appears to be established. The present study concluded that:

$$\begin{aligned}\tau_{BM} &= \frac{\tau_{OAE}}{2}, & f &\lesssim 2 \text{ kHz} \\ \tau_{BM} &= \tau_{OAE}, & f &\gtrsim 4 \text{ kHz}\end{aligned}$$

The boundaries are not precisely known and, in the frequency range 2-4 kHz, the cochlear delay cannot be estimated precisely from OAE measurements. Between 2 and 4 kHz, there seems to be a transition between the two predictable relations. The reasons for having different relations between  $\tau_{OAE}$  and  $\tau_{BM}$  is not completely understood. Several hypotheses have been made in Chap. 5 and the implications for cochlear modelling are discussed in the final section.

### 8.3 Implications and future work

Various models have been proposed to try modelling the mechanics of the cochlea, from simple linear one-dimensional models (Neely and Kim, 1986; Geisler, 1993) to nonlinear and multidimensional ones (Shera et al., 2007; de Boer et al., 2007). An improvement for the linear models would be to introduce a difference of processing between low and high frequencies. In a recent investigation, Shera et al. (2007) compared the fast- and slow-wave models and concluded that OAEs are generated by transverse (slow) travelling waves. Their approach is only tested at high frequencies (above 5 kHz) and it would be of great interest to extend this prediction for lower frequencies. Maybe, as Shera et al. (2007) suggest, the backward travelling wave propagates through the same medium for high and low frequencies. The only difference would then be that, for low frequencies, this backward travel is disturbed by internal reflections. This could explain the difference observed in the present study between high and low frequencies. Using a hydrodynamic model, Nobili et al. (2003) conclude that OAEs are a middle ear artifact. Although surprising and criticized (Shera et al., 2007; Kalluri and Shera, 2007), this study combines a cochlear model with a middle-ear model. It would be interesting in a future work to investigate closely the effect of the middle ear on OAEs and maybe to compensate for any change in phase. As Puria (2003) showed, there is a clear difference between the phase of the middle

ear reverse transfer function at low and high frequencies (Fig. 1(c) of Nobili et al., 2003; Puria, 2003).

In the present work OAEs and ABRs were successfully recorded to estimate cochlear delays. ASSRs were found to confirm the ABR estimates. OAEs are recorded in the ear canal whereas ABRs and ASSRs reflect the cochlear activity at a neuronal level. Electrocochleography could be used as an additional measurement of the cochlear activity. ECochG can be performed invasively, which allows to place an electrode close to the oval window. The cochlear delays hence estimated could be compared with the ones obtained in the OAE experiment and could confirm or disprove the present findings. The experiment should be conducted with the same subjects to take the inter-subject difference into consideration.

Another possible future work could be to calculate a mapping of the cochlea based on the latency-frequency function found here, and to put it in parallel with other cochlear mappings, such as the ones proposed by Greenwood (1990) or Temchin et al. (2005). Since these two predictions were not based on living humans, the cochlear mapping derived from the present work is closer to the "real" functional human cochlea.

Also related to the cochlear mapping, a direct application of the present experiments could be for cochlear implants (CI). Studies about CI have measured the efficiency of the implantation by comparing the frequency-to-place map of the CI with the Greenwood mapping (Baumann and Nobbe, 2006; Boëx et al., 2006; Dorman et al., 2007). The Greenwood mapping was obtained from post-mortem cochleae. Maybe the data collected in this study can be used to create a new standard of cochlear mapping and this would help improving the placement of the electrodes. Hearing impaired persons could also be taken as test subjects for similar experiments (OAE, ABR and ASSR). This could lead to a better understanding of the effect on hearing impairment on cochlear mechanics. The best case would be to test persons with hearing losses at limited frequency regions. No OAE would be expected from these regions but TBOAE above and below such regions could tell more about the cochlear processing, where the OAE are generated and how a local damage of the BM influences the backward travel of OAE, if at all present.

Further investigations are needed to improve the efficiency of the OAE-based chirp. A better compensation of the dispersion occurring at low frequencies could be taken into account in the chirp parameters. More tests are also necessary with normal-hearing subjects to demonstrate the potential of such stimulus. Clinical tests could also be performed with hearing impaired patients to compare the results of Chirp-OAEs with Click-OAEs.



# Bibliography

- Agung, K., Purdy, S. C., Patuzzi, R. B., O'Beirne, G. A., and Newall, P. (2005). Rising-frequency chirps and earphones with an extended high-frequency response enhance the post-auricular muscle response. *Int. J. Audiol.*, 44(11), 631–636.
- ANSI (1973). *American National Standard: specifications for audiometers*. ANSI S3.9-1969.
- Baumann, U. and Nobbe, A. (2006). The cochlear implant electrode-pitch function. *Hear. Res.*, 213(1-2), 34–42.
- Beattie, R. and Ireland, A. (2000). Effects of sample size on the noise floor and distortion product otoacoustic emissions. *Scand. Audiol.*, 29(2), 93–102.
- Beattie, R., Johnson, A., and Garcia, E. (1996). Frequency-specific auditory brainstem responses in adults with normal hearing. *Aust. J. Audiol.*, 18(1), 1–11.
- Beattie, R. C., Aleks, L. A., and Abbott, C. L. (1994). Effects of signal-to-noise ratio on the auditory brainstem response to 0.5 and 2 kHz tone bursts in broadband noise and highpass noise or notch noise. *Scand. Audiol.*, 23(4), 211–223.
- Beattie, R. C. and Boyd, R. L. (1985). Early/middle evoked potentials to tone bursts in quiet, white noise and notched noise. *Audiology*, 24(6), 406–419.
- Beattie, R. C., Moretti, M., and Warren, V. (1984). Effects of rise-fall time, frequency, and intensity on the early/middle evoked response. *J Speech Hear Disord*, 49(2), 114–127.

- Beattie, R. C. and Torre, P. (1997). Effects of rise-fall time and repetition rate on the auditory brainstem response to 0.5 and 1 kHz tone bursts using normal-hearing and hearing-impaired subjects. *Scand. Audiol.*, 26(1), 23–32.
- Bell, S. L., Allen, R., and Lutman, M. E. (2002). An investigation of the use of band-limited chirp stimuli to obtain the auditory brainstem response. *Int. J. Audiol.*, 41(5), 271–278.
- Bendat, J. S. and Piersol, A. G. (1980). *Engineering applications of correlation and spectral analysis* (1st ed.). Wiley-Interscience.
- Berger, H. (1929). Das Elektrenkephalogramm des Menschen. *Die Naturwissenschaften*, 23(8), 121–124.
- Boëx, C., Baud, L., Cosendai, G., Sigrist, A., Kós, M.-I., and Pelizzone, M. (2006). Acoustic to electric pitch comparisons in cochlear implant subjects with residual hearing. *J Assoc Res Otolaryngol*, 7(2), 110–124.
- Brass, D. and Kemp, D. (1991). Time-domain observation of otoacoustic emission during constant tone stimulation. *J. Acoust. Soc. Am.*, 90, 2415–2427.
- Burkard, R. (1991). Human brain-stem auditory evoked responses obtained by cross correlation to trains of clicks, noise bursts, and tone bursts. *J. Acoust. Soc. Am.*, 90(3), 1398–1404.
- Burkard, R. and Hecox, K. (1983). The effect of broadband noise on the human brainstem auditory evoked response. II. Frequency specificity. *J. Acoust. Soc. Am.*, 74(4), 1214–1223.
- Burkard, R. and Secor, C. (2002). Overview of auditory evoked potential. In J. Katz (Ed.), *Handbook of Clinical Audiology* chapter 14, (pp. 233–248). Lippincott, Williams, and Wilkins.
- Cohen, L. T., Rickards, F. W., and Clark, G. M. (1991). A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. *J. Acoust. Soc. Am.*, 90(5), 2467–2479.

- Cone-Wesson, B., Dowell, R. C., Tomlin, D., Rance, G., and Ming, W. J. (2002). The auditory steady-state response: Comparisons with the auditory brainstem response. *J. Am. Acad. Audiol.*, 13(4), 173–187.
- Conijn, E. A., Brocaar, M. P., and van Zanten, G. A. (1990). Frequency specificity of the auditory brainstem response elicited by 1000-Hz filtered clicks. *Audiology*, 29(4), 181–195.
- Dau, T., Wegner, O., Mellert, V., and Kollmeier, B. (2000). Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *J. Acoust. Soc. Am.*, 107(3), 1530–1540.
- Davis, H. and Hirsh, S. K. (1976). The audiometric utility of brain stem responses to low-frequency sounds. *Audiology*, 15(3), 181–195.
- Davis, H., Hirsh, S. K., Turpin, L. L., and Peacock, M. E. (1985). Threshold sensitivity and frequency specificity in auditory brainstem response audiometry. *Audiology*, 24(1), 54–70.
- de Boer, E. (1980). Auditory physics. Physical principles in hearing theory. I. *Phys. Rep.*, 62, 87–174.
- de Boer, E., Nuttall, A. L., and Shera, C. A. (2007). Wave propagation patterns in a “classical” three-dimensional model of the cochlea. *J. Acoust. Soc. Am.*, 121(1), 352–362.
- Dixit, A., Vaney, N., and Tandon, O. P. (2006). Effect of caffeine on central auditory pathways: An evoked potential study. *Hear. Res.*, 220(1-2), 61–66.
- Dobie, R. A. and Wilson, M. J. (1989). Analysis of auditory evoked potentials by magnitude-squared coherence. *Ear Hear*, 10(1), 2–13.
- Dobie, R. A. and Wilson, M. J. (1993). Objective response detection in the frequency domain. *Electroencephalogr. Clin. Neurophysiol.*, 88(6), 516–524.
- Dolphin, W. F. and Mountain, D. C. (1992). The envelope following response: Scalp potentials elicited in the Mongolian gerbil using sinusoidally AM acoustic-signals. *Hear. Res.*, 58(1), 70–78.

- Don, M. and Eggermont, J. J. (1978). Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. *J. Acoust. Soc. Am.*, 63(4), 1084–1092.
- Don, M. and Kwong, B. (2002). Auditory brainstem response: Differential diagnosis. In J. Katz (Ed.), *Handbook of Clinical Audiology* chapter 16, (pp. 274–297). Lippincott, Williams, and Wilkins.
- Don, M., Kwong, B., and Tanaka, C. (2005). A diagnostic test for ménière's disease and cochlear hydrops: Impaired high-pass noise masking of auditory brainstem responses. *Otol Neurotol*, 26(4), 711–722.
- Don, M., Ponton, C. W., Eggermont, J. J., and Kwong, B. (1998). The effects of sensory hearing loss on cochlear filter times estimated from auditory brainstem response latencies. *J. Acoust. Soc. Am.*, 104(4), 2280–2289.
- Don, M., Ponton, C. W., Eggermont, J. J., and Masuda, A. (1993). Gender differences in cochlear response time: An explanation for gender amplitude differences in the unmasked auditory brain-stem response. *J. Acoust. Soc. Am.*, 94(4), 2135–2148.
- Don, M., Ponton, C. W., Eggermont, J. J., and Masuda, A. (1994). Auditory brainstem response peak amplitude variability reflects individual differences in cochlear response times. *J. Acoust. Soc. Am.*, 96(6), 3476–3491.
- Donaldson, G. S. and Ruth, R. A. (1993). Derived band auditory brain-stem response estimates of traveling wave velocity in humans. I: Normal-hearing subjects. *J. Acoust. Soc. Am.*, 93(2), 940–951.
- Dong, W. and Cooper, N. P. (2006). An experimental study into the acousto-mechanical effects of invading the cochlea. *J R Soc Interface*, 3(9), 561–571.
- Dorman, M. F., Spahr, T., Gifford, R., Loiselle, L., McKarns, S., Holden, T., Skinner, M., and Finley, C. (2007). An electric frequency-to-place map for a cochlear implant patient with hearing in the nonimplanted ear. *J Assoc Res Otolaryngol*, 8(2), 234–240.



- Dreisbach, L., Siegel, J., and Chen, W. (1998). Stimulus-frequency otoacoustic emissions measured at low- and high-frequencies in untrained human subjects. In *Assoc. Res. Otolaryngol. Abs.* 21, 349.
- Eggermont, J. and Don, M. (1980). Analysis of the click-evoked brainstem potentials in humans using high-pass noise masking. II. Effect of click intensity. *J. Acoust. Soc. Am.*, 68(6), 1671–1675.
- Eggermont, J. J. (1979a). Compound action potentials: Tuning curves and delay times. *Scand. Audiol., Suppl.* 9, 129–139.
- Eggermont, J. J. (1979b). Narrow-band AP latencies in normal and recruiting human ears. *J. Acoust. Soc. Am.*, 65(2), 463–470.
- Elberling, C. and Don, M. (1984). Quality estimation of averaged auditory brainstem responses. *Scand. Audiol.*, 13(3), 187–197.
- Elberling, C., Don, M., Cebulla, M., and Stürzebecher, E. (2007). Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay. *J. Acoust. Soc. Am.*, 122(5), 2772–2785.
- Elberling, C., Parbo, J., Johnsen, N., and Bagi, P. (1985). Evoked acoustic emission: Clinical application. *Acta Oto-Laryngol.*, 77–85. Suppl. 421.
- Fausti, S. A., Olson, D. J., Frey, R. H., Henry, J. A., and Schaffer, H. I. (1993). High-frequency tone burst-evoked ABR latency-intensity functions. *Scand. Audiol.*, 22(1), 25–33.
- Fobel, O. and Dau, T. (2004). Searching for the optimal stimulus eliciting auditory brainstem responses in humans. *J. Acoust. Soc. Am.*, 116(4), 2213–2222.
- Foxe, J. J. and Stapells, D. R. (1993). Normal infant and adult auditory brain-stem responses to bone-conducted tones. *Audiology*, 32(2), 95–109.
- Galampos, R., Makeig, S., and Talmachoff, P. J. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proc. Natl. Acad. Sci. USA*, 78(4), 2643–2647.

- Geisler, C. D. (1993). A realizable cochlear model using feedback from motile outer hair cells. *Hear. Res.*, 68(2), 253–262.
- Gold, T. (1948). Hearing. II. The physical basis of the action of the cochlea. *Proc. R. Soc. London, Ser. B*, 135(881), 492–498.
- Goodman, S., Withnell, R., de Boer, E., Lilly, D., and Nuttall, A. (2004). Cochlear delays measured with amplitude-modulated tone-burst-evoked OAEs. *Hear. Res.*, 188(1-2), 57–69.
- Gorga, M., Kaminski, J., Beauchaine, K., and Jesteadt, W. (1988). Auditory brainstem responses to tone bursts in normally hearing subjects. *J. Speech Hear. Res.*, 31(1), 87–97.
- Gorga, M., Neely, S., Bergman, B., Beauchaine, K., Kaminski, J., and Liu, Z. (1994). Towards understanding the limits of distortion product otoacoustic emission measurements. *J. Acoust. Soc. Am.*, 96(3), 1494–1500.
- Gorga, M. P., Neely, S. T., and Dorn, P. A. (2000). Distortion product otoacoustic emission test performance when both  $2f_1 - f_2$  and  $2f_2 - f_1$  are used to predict auditory status. *Ear Hear*, 20(4), 345–362.
- Grandori, F. (1985). Nonlinear phenomena in click- and tone-burst-evoked otoacoustic emissions from human ears. *Audiology*, (24), 71–80.
- Greenwood, D. (1990). A cochlear frequency-position function for several species-29 years later. *J. Acoust. Soc. Am.*, 87(6), 2592–2605.
- Gummer, A. (Ed.). (2003). *Biophysics of the Cochlea: from Molecules to Models*. World Scientific.
- Hall, J. W. (2006). *New Handbook for Auditory Evoked Responses*. Allyn & Bacon.
- Hall, J. W., Bull, J. M., and Cronau, L. H. (1988). Hypo- and hyperthermia in clinical auditory brain stem response measurement: Two case reports. *Ear Hear*, 9(3), 137–143.

- Harte, J. M. and Elliott, S. J. (2005). Using the short-time correlation coefficient to compare transient- and derived, noise-evoked otoacoustic emission temporal waveforms. *J. Acoust. Soc. Am.*, 117(5), 2989–2998.
- He, W., Nuttall, A. L., and Ren, T. (2007). Two-tone distortion at different longitudinal locations on the basilar membrane. *Hear. Res.*, 228(1-2), 112–122.
- Hecox, K. and Deegan, D. (1983). Rise-fall time effects on the brainstem auditory evoked response: Mechanisms. *J. Acoust. Soc. Am.*, 73(6), 2109–2116.
- Hecox, K., Squires, N., and Galambos, R. (1976). Brainstem auditory evoked responses in man. I. Effect of stimulus rise–fall time and duration. *J. Acoust. Soc. Am.*, 60(5), 1187–1192.
- Heil, P. and Irvine, D. R. (1997). First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol*, 78(5), 2438–2454.
- Herdman, A. T., Lins, O., Roon, P. V., Stapells, D. R., Scherg, M., and Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topogr.*, 15(2), 69–86.
- Hoth, S. and Weber, F. (2001). The latency of evoked otoacoustic emissions: Its relation to hearing loss and auditory evoked potentials. *Scand. Audiol.*, 30(3), 173–183.
- IEC (2007). *60645-3 Audiometers. Part 3: Auditory test signals of short duration for audiometric and neuro-otological purposes*. (2.0 ed.). [www.iec.ch](http://www.iec.ch): International Electrotechnical Commission.
- ISO (1994). *389-2. Reference zero for the calibration of audiometric equipment: Part 2. Reference equivalent threshold sound pressure levels for pure tones and insert earphones*. [www.iso.org](http://www.iso.org): International Organization for Standardization.
- Jasper, H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.*, 10, 371–375.

- Jedrzejczak, W., Blinowska, K., and Konopka, W. (2005). Time-frequency analysis of transiently evoked otoacoustic emissions of subjects exposed to noise. *Hear. Res.*, 205(1-2), 249–255.
- Jewett, D. L., Romano, M. N., and Williston, J. S. (1970). Human auditory evoked potentials: Possible brain stem components detected on the scalp. *Science*, 167(924), 1517–1518.
- Jewett, D. L. and Williston, J. (1971). Auditory-evoked far fields averaged from scalp of humans. *Brain*, 94, 681–696.
- John, M. S., Dimitrijevic, A., van Roon, P., and Picton, T. W. (2001). Multiple auditory steady-state responses to AM and FM stimuli. *Audiol. Neurootol.*, 6(1), 12–27.
- John, M. S., Lins, O. G., Boucher, B. L., and Picton, T. W. (1998). Multiple auditory steady-state responses (MASTER): Stimulus and recording parameters. *Audiology*, 37(2), 59–82.
- John, M. S. and Picton, T. W. (2000). Human auditory steady-state responses to amplitude-modulated tones: Phase and latency measurements. *Hear. Res.*, 141(1-2), 57–79.
- Johnsen, N. and Elberling, C. (1982). Evoked acoustic emissions from the human ear. II. Normative data in young adults and influence of posture. *Scand. Audiol.*, 11(1), 69–77.
- Junius, D. and Dau, T. (2005). Influence of cochlear traveling wave and neural adaptation on auditory brainstem responses. *Hear. Res.*, 205(1-2), 53–67.
- 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. (2000). Nonlinear temporal interactions in click-evoked otoacoustic emissions. II. Experimental data. *Hear. Res.*, 146(1), 101–120.
- Katz, J. (2001). *Handbook of Clinical Audiology*. Lippincott, Williams, and Wilkins.

- Keefe, D. (1998). Double-evoked otoacoustic emissions. I. Measurement theory and nonlinear coherence. *J. Acoust. Soc. Am.*, 103(6), 3489–98.
- Kemp, D. (2003). Twenty five years of OAEs. *ENT News*, 12(5), 47–52.
- Kemp, D., Bray, P., Alexander, L., and Brown, A. (1986). Acoustic emission cochleography-practical aspects. *Scand. Audiol. Suppl.*, 1(25), 71–95.
- Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *J. Acoust. Soc. Am.*, 64(5), 1386–1391.
- Kemp, D. T. (1986). Otoacoustic emissions, travelling waves and cochlear mechanisms. *Hear. Res.*, 22, 95–104.
- Kemp, D. T. (2002). Exploring cochlear status with otoacoustic emissions: The potential for new clinical applications. In M. S. Robinette and T. J. Glatke (Eds.), *Otoacoustic Emissions: Clinical Applications* chapter 1, (pp. 1–47). Thieme.
- Kemp, D. T. and Chum, R. (1980). Properties of the generator of stimulated acoustic emissions. *Hear. Res.*, 2, 213–232.
- Kiang, N. Y. (1975). Stimulus representation in the discharge patterns of auditory neurons. In E. L. Eagles (Ed.), *The Nervous System. Volume 3: Human Communication and Its Disorders* (pp. 81–96). Raven Press.
- Kileny, P. (1981). The frequency specificity of tone-pip evoked auditory brain stem responses. *Ear Hear*, 2(6), 270–275.
- Killan, E. and Kapadia, S. (2006). Simultaneous suppression of tone burst-evoked otoacoustic emissions - Effect of level and presentation paradigm. *Hear. Res.*, 212(1-2), 65–73.
- Kim, D. O. and Molnar, C. E. (1979). A population study of cochlear nerve fibers: Comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones. *J Neurophysiol*, 42(1 Pt 1), 16–30.
- Kodera, K., Yamane, H., Yamada, O., and Suzuki, J. I. (1977a). Brain stem response audiometry at speech frequencies. *Audiology*, 16(6), 469–479.

- Kodera, K., Yamane, H., Yamada, O., and Suzuki, J. I. (1977b). The effect of onset, offset and rise-decay times of tone bursts on brain stem response. *Scand. Audiol.*, 6(4), 205–210.
- Konrad-Martin, D. and Keefe, D. H. (2003). Time-frequency analyses of transient-evoked stimulus-frequency and distortion-product otoacoustic emissions: Testing cochlear model predictions. *J. Acoust. Soc. Am.*, 114(4), 2021–2043.
- Kuwada, S., Anderson, J. S., Batra, R., Fitzpatrick, D. C., Teissier, N., and D'Angelo, W. R. (2002). Sources of the scalp-recorded amplitude-modulation following response. *J. Am. Acad. Audiol.*, 13(4), 188–204.
- Kuwada, S., Batra, R., and Maher, V. L. (1986). Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. *Hear. Res.*, 21(2), 179–192.
- Levi, E. C., Folsom, R. C., and Dobie, R. A. (1993). Amplitude-modulation following response (AMFR): Effects of modulation rate, carrier frequency, age, and state. *Hear. Res.*, 68(1), 42–52.
- Lins, O. G. and Picton, T. W. (1995). Auditory steady-state responses to multiple simultaneous stimuli. *Electroencephalogr. Clin. Neurophysiol.*, 96(5), 420–432.
- Long, G. and Talmadge, C. (1997). Spontaneous otoacoustic emission frequency is modulated by heartbeat. *J. Acoust. Soc. Am.*, 102(5), 2831–2848.
- Lucertini, M., Moleti, A., and Sisto, R. (2002). On the detection of early cochlear damage by otoacoustic emission analysis. *J. Acoust. Soc. Am.*, 111(2), 972–978.
- Maurizi, M., Paludetti, G., Ottaviani, F., and Rosignoli, M. (1984). Auditory brainstem responses to middle- and low-frequency tone pips. *Audiology*, 23(1), 75–84.
- Moleti, A. and Sisto, R. (2008). Comparison between otoacoustic and auditory brainstem response latencies supports slow backward propagation of otoacoustic emissions. *J. Acoust. Soc. Am.*, 123(3), 1495–1503.

- Moleti, A., Sisto, R., Tognola, G., Parazzini, M., Ravazzani, P., and Grandori, F. (2005). Otoacoustic emission latency, cochlear tuning, and hearing functionality in neonates. *J. Acoust. Soc. Am.*, 118(3), 1576–1584.
- Møller, A. R. and Jannetta, P. J. (1983). Interpretation of brainstem auditory evoked-potentials: Results from intracranial recordings in humans. *Scand. Audiol.*, 12(2), 125–133.
- Moore, B. C. (2003). *Psychology of Hearing* (Fifth ed.). Academic Press.
- Murray, J. G., Cohn, E. S., Harker, L. A., and Gorga, M. P. (1998). Tone burst auditory brain stem response latency estimates of cochlear travel time in Meniere's disease, cochlear hearing loss, and normal ears. *Am. J. Otol.*, 19(6), 854–859.
- Narayan, S. S. (1991). *Comparison of latencies of NI and transient evoked otoacoustic emissions: An evaluation of reverse travel in the cochlea*. PhD thesis, Purdue University.
- Neely, S., Norton, S., Gorga, M., and Jesteadt, W. (1988). Latency of auditory brainstem responses and otoacoustic emissions using tone-burst stimuli. *J. Acoust. Soc. Am.*, 83(2), 652–656.
- Neely, S. T. and Kim, D. O. (1986). A model for active elements in cochlear biomechanics. *J. Acoust. Soc. Am.*, 79(5), 1472–1480.
- Nelson, D. and Zhou, J. (1996). Slopes of distortion-product otoacoustic emission growth curves corrected for noise-floor levels. *J. Acoust. Soc. Am.*, 99(1), 468–474.
- Nobili, R., Vetesnik, A., Turicchia, L., and Mammano, F. (2003). Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model. *J. Assoc. Res. Otolaryngol.*, 4(4), 478–494.
- Norton, S. and Neely, S. (1987). Tone-burst-evoked otoacoustic emissions from normal-hearing subjects. *J. Acoust. Soc. Am.*, 81(6), 1860–1872.
- Oates, P. and Stapells, D. R. (1997). Frequency specificity of the human auditory brainstem and middle latency responses to brief tones. I. High-pass noise masking. *J. Acoust. Soc. Am.*, 102(6), 3597–3608.

- Oppenheim, A., Schafer, R., and Buck, J. R. (1998). *Discrete-Time Signal Processing*. Prentice-Hall.
- Patuzzi, R. (1996). Cochlear micromechanics and macromechanics. In P. Dallos, A. Popper, and R. Fay (Eds.), *The Cochlea* chapter 4, (pp. 186–257). Springer-Verlag.
- Picton, T. W., Dimitrijevic, A., John, M. S., and van Roon, P. (2001). The use of phase in the detection of auditory steady-state responses. *Clin. Neurophysiol.*, 112(9), 1698–1711.
- Picton, T. W., Hillyard, S. A., Krausz, H. I., and Galambos, R. (1974). Human auditory evoked-potentials. I: Evaluation of components. *Electroencephalogr. Clin. Neurophysiol.*, 36(2), 179–190.
- Probst, R., Coats, A., Martin, G., and Lonsbury-Martin, B. (1986). Spontaneous, click-, and toneburst-evoked otoacoustic emissions from normal ears. *Hear. Res.*, 21(3), 261–275.
- Probst, R., Lonsbury-Martin, B., and Martin, G. (1991). A review of otoacoustic emissions. *J. Acoust. Soc. Am.*, 89(5), 2027–2067.
- Purcell, D. W., John, M. S., and Picton, T. W. (2003). Concurrent measurement of distortion product otoacoustic emissions and auditory steady state evoked potentials. *Hear. Res.*, 176(1-2), 128–141.
- Purcell, D. W., Kunov, H., and Cleghorn, W. (2003). Estimating bone conduction transfer functions using otoacoustic emissions. *J. Acoust. Soc. Am.*, 114(2), 907–918.
- Purcell, D. W., Van Roon, P., Sasha John, M., and Picton, T. W. (2006). Simultaneous latency estimations for distortion product otoacoustic emissions and envelope following responses. *J. Acoust. Soc. Am.*, 119(5), 2869–2880.
- Puria, S. (2003). Measurements of human middle ear forward and reverse acoustics: implications for otoacoustic emissions. *J. Acoust. Soc. Am.*, 113(5), 2773–2789.



- Rance, G., Tomlin, D., and Rickards, F. W. (2006). Comparison of auditory steady-state responses and tone-burst auditory brainstem responses in normal babies. *Ear Hear*, 27(6), 751–762.
- Ren, T. (2004). Reverse propagation of sound in the gerbil cochlea. *Nat. Neurosci.*, 7(4), 333–334.
- Rickards, F. W., Tan, L. E., Cohen, L. T., Wilson, O. J., Drew, J. H., and Clark, G. M. (1994). Auditory steady-state evoked potential in newborns. *Br. J. Audiol.*, 28(6), 327–337.
- Riedel, H., Granzow, M., and Kollmeier, B. (2001). Single-sweep-based methods to improve the quality of auditory brain stem responses Part II: Averaging methods. *Z. Audiol.*, 40(2), 62–85.
- Robinette, M. S. and Glatke, T. J. (2001). *Otoacoustic Emissions* (2nd ed.). Thieme Medical Publishers.
- Robles, L. and Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiol. Rev.*, 81(3), 1305–1352.
- Rodriguez, R., Picton, T., Linden, D., Hamel, G., and Laframboise, G. (1986). Human auditory steady state responses: Effects of intensity and frequency. *Ear Hear*, 7(5), 300–313.
- Ruggero, M. (2007). Personal communication.
- Ruggero, M. A. (2004). Comparison of group delays of  $2f_1 - f_2$  distortion product otoacoustic emissions and cochlear travel. *ARLO*, 4(5), 143–147. Acoustics Research Letters Online.
- Ruggero, M. A., Rich, N. C., Recio, A., 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.
- Ruggero, M. A. and Temchin, A. N. (2007). Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods. *J Assoc Res Otolaryngol*, 8(2), 153–166.

- Rutten, W. (1980). Evoked acoustic emission from within normal and abnormal human ears: Comparison with audiometric and electrocochleographic findings. *Hear. Res.*, 2, 263–271.
- Schairer, K. S., Ellison, J. C., 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.
- Schairer, K. S., Fitzpatrick, D., and Keefe, D. H. (2003). Input-output functions for stimulus-frequency otoacoustic emissions in normal-hearing adult ears. *J. Acoust. Soc. Am.*, 114(2), 944–966.
- Schönweiler, R., Neumann, A., and Ptok, M. (2005). Frequency specific auditory evoked responses. Experiments on stimulus polarity, sweep frequency, stimulus duration, notched-noise masking level, and threshold estimation in volunteers with normal hearing. *HNO*, 53(11), 983–994.
- Schoonhoven, R., Boden, C. J. R., Verbunt, J. P. A., and de Munck, J. C. (2003). A whole head MEG study of the amplitude-modulation-following response: Phase coherence, group delay and dipole source analysis. *Clin. Neurophysiol.*, 114(11), 2096–2106.
- Schoonhoven, R., Prijs, V., and Schneider, S. (2001). DPOAE group delays versus electrophysiological measures of cochlear delay in normal human ears. *J. Acoust. Soc. Am.*, 109(4), 1503–1512.
- Serbetçioğlu, M. B. and Parker, D. J. (1999). Measures of cochlear travelling wave delay in humans: I. Comparison of three techniques in subjects with normal hearing. *Acta Oto-Laryngol.*, 119(5), 537–543.
- 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. (2003). Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning. *J. Acoust. Soc. Am.*, 113(5), 2762–2772.

- Shera, C. A., Tubis, A., Talmadge, C. L., de Boer, E., Fahey, P. F., and Guinan, J. J. (2007). Allen-Fahey and related experiments support the predominance of cochlear slow-wave otoacoustic emissions. *J. Acoust. Soc. Am.*, 121(3), 1564–1575.
- Siegel, J. H., Cerka, A. J., Recio-Spinoso, A., Temchin, A. N., Van Dijk, P., and Ruggero, M. A. (2005). Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering. *J. Acoust. Soc. Am.*, 118(4), 2434–2443.
- Sininger, Y. and Cone-Wesson, B. (2002). Threshold prediction using auditory brainstem response and steady-state evoked potentials with infants and young children. In J. Katz (Ed.), *Handbook of Clinical Audiology* chapter 17, (pp. 298–322). Lippincott, Williams and Wilkins.
- Sisto, R. and Moleti, A. (2002). On the frequency dependence of the otoacoustic emission latency in hypoacoustic and normal ears. *J. Acoust. Soc. Am.*, 111(1), 297–308.
- Sisto, R. and Moleti, A. (2007). Transient evoked otoacoustic emission latency and cochlear tuning at different stimulus levels. *J. Acoust. Soc. Am.*, 122(4), 2183–2190.
- Squires, K. C., Chu, N. S., and Starr, A. (1978). Acute effects of alcohol on auditory brainstem potentials in humans. *Science*, 201(4351), 174–176.
- Stapells, D. R. (1994). Low-frequency hearing and the auditory brainstem response. *Am. J. Audiol.*, 7, 11–13.
- Stapells, D. R. and Picton, T. W. (1981). Technical aspects of brainstem evoked potential audiometry using tones. *Ear Hear*, 2(1), 20–29.
- Stover, L. and Norton, S. (1993). The effects of aging on otoacoustic emissions. *J. Acoust. Soc. Am.*, 94(5), 2670–2681.
- Strube, H. W. (1989). Evoked otoacoustic emissions as cochlear Bragg reflections. *Hear. Res.*, 38(1-2), 35–45.

- Stürzebecher, E., Cebulla, M., Elberling, C., and Berger, T. (2006). New efficient stimuli for evoking frequency-specific auditory steady-state responses. *J. Am. Acad. Audiol.*, 17(6), 448–461.
- Suzuki, T., Hirai, Y., and Horiuchi, K. (1977). Auditory brain stem responses to pure tone stimuli. *Scand. Audiol.*, 6(1), 51–56.
- Suzuki, T. and Horiuchi, K. (1977). Effect of high-pass filter on auditory brain stem responses to tone pips. *Scand. Audiol.*, 6(3), 123–126.
- Suzuki, T. and Horiuchi, K. (1981). Rise time of pure-tone stimuli in brain stem response audiometry. *Audiology*, 20(2), 101–112.
- Temchin, A. N., Recio-Spinoso, A., van Dijk, P., and Ruggero, M. A. (2005). Wiener kernels of chinchilla auditory-nerve fibers: Verification using responses to tones, clicks, and noise and comparison with basilar-membrane vibrations. *J. Neurophysiol.*, 93(6), 3635–3648.
- Terkildsen, K., Osterhammel, P., and in't Veld, F. H. (1973). Electrocochleography with a far field technique. *Scand. Audiol.*, 2(3), 141–148.
- Thornton, A., Lineton, B., Baker, V., and Slaven, A. (2006). Nonlinear properties of otoacoustic emissions in normal and impaired hearing. *Hear. Res.*, 219(1-2), 56–65.
- Tognola, G., Grandori, F., and Ravazzani, P. (1997). Time-frequency distributions of click-evoked otoacoustic emissions. *Hear. Res.*, 106(1-2), 112–122.
- van der Reijden, C. S., Mens, L. H. M., and Snik, A. F. M. (2006). Frequency-specific objective audiometry: Tone-evoked brainstem responses and steady-state responses to 40 Hz and 90 Hz amplitude modulated stimuli. *Int. J. Audiol.*, 45(1), 40–45.
- Wegner, O. and Dau, T. (2002). Frequency specificity of chirp-evoked auditory brainstem responses. *J. Acoust. Soc. Am.*, 111(3), 1318–1329.
- Whitehead, M., Stagner, B., Martin, G., and Lonsbury-Martin, B. (1996). Visualization of the onset of distortion product otoacoustic emissions, and measurement of their latency. *J. Acoust. Soc. Am.*, 100(3), 1663–1679.

- Wilson, J. P. (1980a). Evidence for a cochlear origin for acoustic re-emission, threshold fine structure and tonal tinnitus. *Hear. Res.*, 2, 233–252.
- Wilson, J. P. (1980b). Model for cochlear echoes and tinnitus based on an observed electrical correlate. *Hear. Res.*, 2(3-4), 527–532.
- Wit, H. and Ritsma, R. (1980). Evoked acoustical responses from the human ear: Some experimental results. *Hear. Res.*, 2, 253–261.
- Zweig, G. and Shera, C. (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. *J. Acoust. Soc. Am.*, 98(4), 2018–2047.
- Zwicker, E. and Lumer, G. (1985). Evaluating traveling wave characteristics in man by an active nonlinear cochlear preprocessing model. In Allen, J., Hall, J., Hubbard, A., Neely, S., and Tubis, A. (Eds.), *Peripheral Auditory Mechanisms*, (pp. 250–257). Springer.



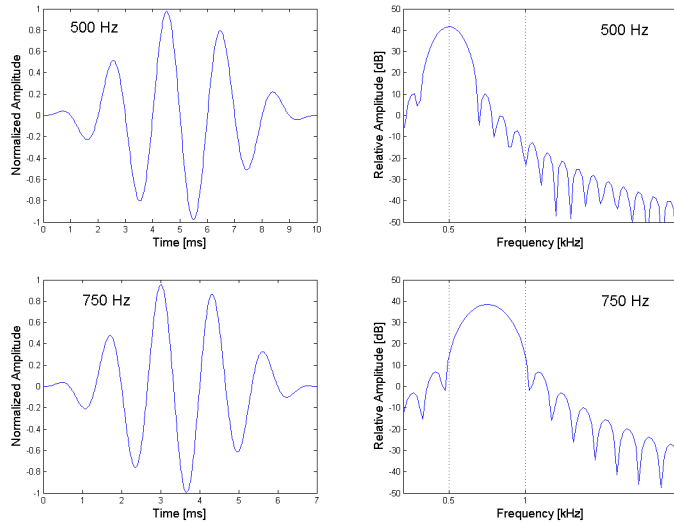
---

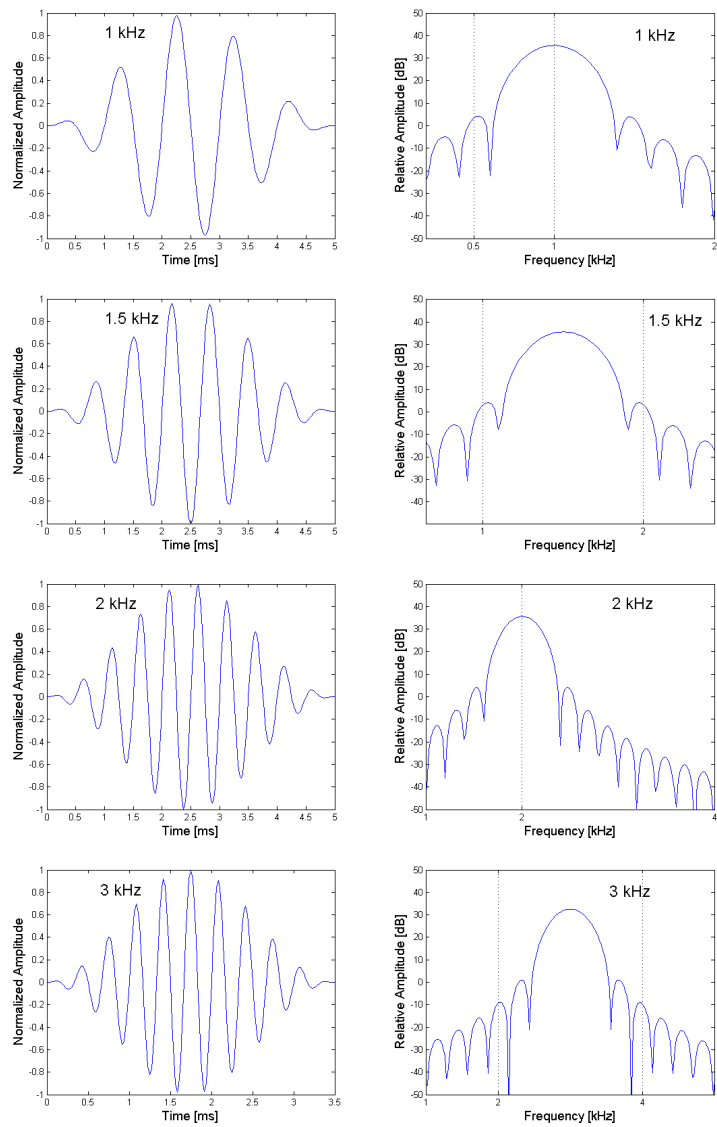
## Appendix A

# Tone bursts: time and frequency plots

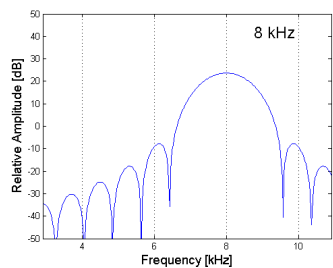
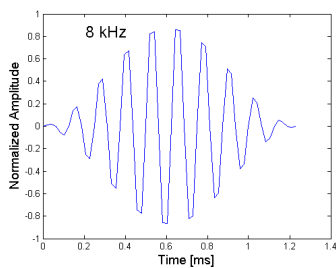
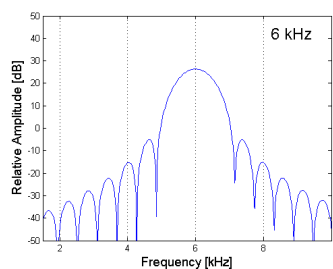
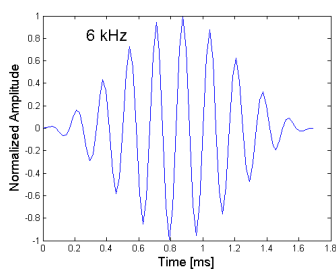
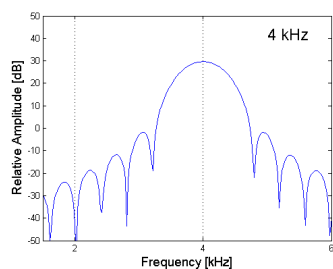
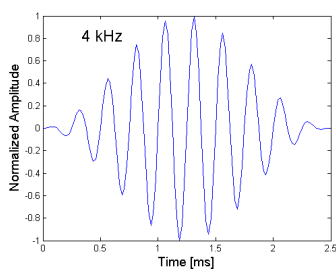
---

In this appendix, the stimuli used in the OAE and ABR experiments are plotted in the time and frequency domain. The length of these tone bursts are a compromise between a narrow spread in frequency and a short stimulus in time. It was tried to reach a similar number of cycles for all frequencies so that the basilar membrane is uniformly excited for all frequencies.











### Stimuli calibration

---

Some calibration is necessary in order to have the correct sound pressure at the ear of the subject. The stimuli used were of very short duration and an equivalent sound pressure level is found to compare with the sound pressure level of longer stimuli. The stimuli were played in the insert earphone ER-2 connected to the ear simulator (B&K 4157). The output of this coupler was fed into the sound level meter (B&K 2607) which was also connected to an oscilloscope, see Fig. B.1. The voltage from the ear simulator could be read both by the sound level meter as RMS values and by the oscilloscope as voltage (peak values).

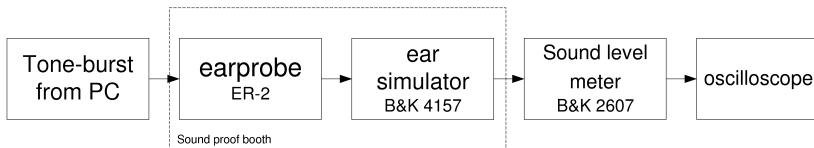


Figure B.1: Setup to find the peak-equivalent SPL of the different tone bursts.

The first step was to play the stimuli and measure their peak-to-peak value on the oscilloscope. Then, a pure-tone stimulus with similar frequency was created and played, its amplitude was adjusted so that its peak-to-peak value was the same as the short-duration stimulus, see figure B.2. The SPL value of this latter was then read on the sound level meter, this gives the peak-equivalent sound pressure level (peSPL) for the tone-burst stimulus. Hence, all references to tone-burst sound pressure levels are indicated in peak-equivalent SPL (pe SPL).

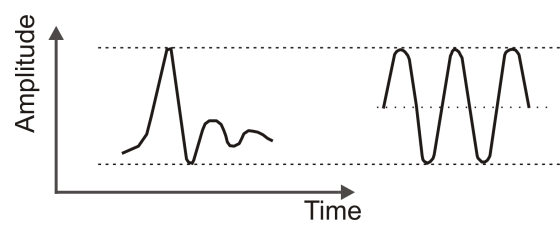


Figure B.2: Equivalence in amplitude between a brief tone (left) and a long duration sinusoid.

## Mathematical tools

---

### C.1 Hilbert transform

In the first experiment (chapter 3), the Hilbert transform was used to obtain the envelope of the recorded signal (containing the OAE response). The envelope of a signal is calculated as the absolute amplitude of the corresponding analytical signal. If  $s(t)$  denote the signal then the corresponding analytical signal is given by  $\hat{s}(t) = s(t) + j\tilde{s}(t)$ , where  $\tilde{s}(t)$  is the Hilbert transform of the original signal and  $j$  is the imaginary unit.

This can also be written  $\hat{s}(t) = A(t)e^{j\phi(t)}$  where  $A(t)$  is the amplitude of the analytic signal, i.e.  $A(t) = |\hat{s}(t)| = \sqrt{s(t)^2 + \tilde{s}(t)^2}$  and  $\phi(t)$  is the phase of  $\hat{s}(t)$ .

To tell a bit more about the Hilbert transform, it is defined as the convolution of  $s(t)$  by  $\frac{1}{\pi t}$ , i.e.  $\tilde{s}(t) = s(t) * \frac{1}{\pi t}$ . The fast Fourier transform (FFT) of the Hilbert transform becomes  $\tilde{S}(f) = S(f) \cdot FFT(\frac{1}{\pi t})$ . This leads to the one sided spectrum of the analytical signal

$$\hat{S}(f) = \begin{cases} 2S(f) & f > 0 \\ S(0) & f = 0 \\ 0 & f < 0 \end{cases}$$

which is simply the right hand side of the original signal with a doubled amplitude. The factor 2 is explain by a need to maintain the same amount of energy in the signal.

The Hilbert transform is named after David Hilbert (1873-1943) who was chairman of mathematics at the university of Göttingen. More explanations about the Hilbert transform can be found in Oppenheim et al. (1998).

## C.2 The least-squares fit time-domain filtering

The least-squares fit (LSF) algorithm is used to detect the TBOAE, as described in section 3.2. The idea of the algorithm is to help detecting a signal (OAE for instance) embedded in background noise. The LSF algorithm fits a signal to the OAE recording, assuming that it can be modeled as

$$y = a \cos[(\omega + \Delta_\omega)t] + b \sin[(\omega + \Delta_\omega)t]$$

where  $\omega$  is the supposed frequency of the signal (here TB centre frequency),  $\Delta_\omega$  is a small offset around  $\omega$  which allows to detect the signal despite a slight change in frequency (it occurs for SOAE (Probst et al., 1991)). It is assumed that  $\Delta_\omega$  is small ( $\Delta_\omega \ll 1$ ) which allows to linearize  $y$ .  $a$  and  $b$  are the unknown amplitudes of the signal components.

The LSF algorithm finds the parameters,  $a$ ,  $b$  and  $\Delta_\omega$  and return the amplitude  $A = \sqrt{a^2 + b^2}$  of the signal. The algorithm proved to be more precise than a FFT analysis. It is based on a study by Long and Talmadge (1997), who originally used this algorithm to detect SOAE. The Matlab script used in the present study to retrieve the OAE amplitude was kindly provided by Manfred Mauermann.

## C.3 Short-time correlation coefficient

The short-time correlation coefficient is calculated to know how two signals evolve compare to one another. It is given by

$$STCC(t) = \frac{\sum[w(t)x(t)y(t)]}{\sqrt{\sum[w(t)x(t)]^2 \cdot \sum[w(t)y(t)]^2}}$$

where  $w(t)$  is a moving window function. The correlation between  $x(t)$  and  $y(t)$  is calculated by shifting this window along the time axis. The window used here is defined as

$$\begin{aligned} w(t) &= 1, \text{ for } t_1 \leq t \leq t_2, \\ &= 0, \text{ for } t \leq t_1 \text{ and } t \geq t_2 \end{aligned}$$

The length of the window is 64 samples (1.3 ms at 48 kHz), this is a compromise found since a too long window would give imprecise results (due to smear out). Besides, to improve the quality of the correlation an overlap of 75% was performed. The summation are made over the entire length of the signals.

The closer STCC is to 1, the more  $x(t)$  and  $y(t)$  are correlated. In the present study, the STCC was calculated between the different runs of TBOAE.

The reader is referred to Harte and Elliott (2005) for more explanation on the STCC method.

## C.4 Pearson product moment correlation coefficient

The aim of this statistical tool is to measure the relationship between the two variables  $x$  and  $y$ ,  $x$  being the SNR value across frequencies and  $y$  the standard deviation of the TBOAE latency estimates across frequencies, see section 3.3.5. A correlation coefficient can be calculated to know how the two variables evolve relatively to each other. The Pearson Product Moment Correlation Coefficient (more simply Correlation Coefficient) is defined as follow:

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

where  $r_{xy}$  is the correlation coefficient,  $s_{xy}$  the covariance,  $s_x$  the standard deviation of  $x$  and  $s_y$  the standard deviation of  $y$ . This can also be written

$$r_{xy} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

where  $x_i$  and  $y_i$  are the values for specific frequencies and  $\bar{x}$  and  $\bar{y}$  are their mean values. The correlation coefficient ranges from -1 to +1. If the value is close to -1 or +1 this indicates a strong linear relationship between the two variables  $x$  and  $y$ . The closer the coefficient is to zero the weaker the correlation is.

In the present study, this correlation coefficient was calculated to see if there was any relation between the SNR and the inter-subject variability. The value found in this case was  $r = 0.04$ , meaning that there is no correlation between SNR and inter-subject variability. It is therefore not believed that SNR has an effect on the OAE latency estimates.

Karl Pearson (1857-1936) was a British statistician who founded the statistics department at University College London.



---

## Appendix D

### TBOAE and TBABR latencies

---

Frequencies [kHz]		0.5	0.75	1	1.5	2	3	4	6	8
Subj 3	Avg	11.1	10.2	8.7	8.2	8.0	7.6	7.3	6.9	6.4
	Std	1.1	0.5	0.6	0.2	0.1	0.1	0.1	0.1	0.1
Subj 4	Avg	10.3			8.7	8.9	8.2	7.9	7.1	6.7
	Std	0			0.8	0.6	0.4	0.2	0.1	0.1
Subj 5	Avg			9.1	8.6	8.3	7.6	7.2	6.9	6.6
	Std			0.7	0.1	0.1	0.2	0.3	0.1	0.1
Subj 6	Avg		11.4	11.3	9.4	8.7	7.8	8	7.3	6.7
	Std		0	0	0.7	0.1	0.1	0.4	0.5	0.2
Subj 7	Avg		12.2	10.3	9.8	9.4	8.6	8.3	7.5	6.7
	Std		0	0.5	0.1	0	0.1	0.2	0.1	0.6
Subj 9	Avg	11.5	10.4	10.1	9.1	8.8	7.9	7.5	6.8	6.6
	Std	0.9	0.5	0.3	0.1	0.1	0.4	0.2	0.1	0.1
Subj 10	Avg	12.8	10.9	10.9	9.9	9.5	8.7	8.5	8.2	7.3
	Std	0.6	0.2	0.1	0.6	0.1	0.4	0.1	0.1	0.3
Subj 11	Avg	11.9	10.4	10.1	9.5	9.4	8.9	8.2	7.6	6.9
	Std	0	0.1	0.3	0.1	0.2	0.2	0.3	0.1	0.1
Subj 12	Avg	12	11.3	11.1	9.4	9.3	8.1	7.8	7.3	6.9
	Std	0.3	0.1	0.1	0	0.1	0.1	0.2	0.2	0.1
Subj 15	Avg	11.6	10.4	10.1	9.1	9.2	8.5	8.2	7.5	7.3
	Std	1.1	0.1	0.1	0.7	0.1	0.1	0.2	0.2	0.1
Subj 16	Avg	11.5	10.2	8.7	9.1	9.1	8.4	8.1	7.6	7.4
	Std	0.2	0.2	0	0.8	0.1	0.3	0.1	0.1	0.1

Table D.1: wave-V latency estimated in the 11 subjects participating in the ABR experiment. Blanks indicate that no values could be found.

Frequencies [kHz]		0.5	0.75	1	1.5	2	3	4	6	8
Subj 1	Avg		13.8		9.4	8.7	5.7	4.5	2.9	2.7
	Std		0.4		0.2	0.3	0.4	0.1	0.6	0.5
Subj 2	Avg	14.4	12.2	10.5	7.4	6.7		4.0	3.2	2.3
	Std	0.5	0.5	0.3	0.1	0.4		0.1	0.9	0.0
Subj 3	Avg	14.5	11.4	8.9	7.9	6.8	5.7	3.9	3	2.1
	Std	0.1	0.7	0.0	0.1	0.1	0.7	0.1	0.7	0.2
Subj 4	Avg		10.7	7.2	6.6	7.0	5.5	3.8	3.0	2.2
	Std		0.0	0.2	0.1	0.0	0.2	0.4	0.8	0.2
Subj 5	Avg	16.5	11.3	8.1	7.4	7.1	5.0	3.9	3.1	2.5
	Std	0.8	0.4	0.3	0.5	0.4	0.2	0.3	0.6	0.2
Subj 6	Avg	14.3	11.0	8.0	8.0	6.9	5.4	4.0	2.7	2.5
	Std	0.3	0.3	0.2	0.4	0.2	0.4	0.1	0.0	0.1
Subj 7	Avg	16.2	11.4	9.6	7.9	6.5	5.1	3.9	2.7	2.5
	Std	0.5	0.1	0.1	0.1	0.0	0.3	0.1	0.1	0.2
Subj 8	Avg	15.6	13.7	11.2	7.8	7.1	5.9	3.9	3.0	2.3
	Std	0.8	0.5	0.8	0.6	0.4	0.1	0.2	0.8	0.1
Subj 9	Avg	14.7	13.8	9.1	7.1	8.0	5.9	4.9	2.6	2.6
	Std	0.2	0.0	0.4	0.1	0.3	0.1	0.3	0.1	0.0
Subj 10	Avg	16.2	12.2	8.9	8.8	8.6	6.7	4.9	3.9	2.7
	Std	1.0	0.3	0.3	0.7	0.4	0.1	0.2	0.2	0.4
Subj 11	Avg	16.2	15.1	11.4	9.3	7.1	5.7	4.4	2.7	2.7
	Std	0.0	0.8	0.5	0.3	0.8	0.2	0.5	0.2	0.1
Subj 12	Avg		10.7	11.5	8.3	7.3	5.0	3.4	2.7	3.0
	Std		0.0	0.7	0.2	0.8	0.2	0.1	0.2	0.0
Subj 13	Avg	15.2	10.8	8.2	7.4	7.5	4.9	3.7	2.6	2.3
	Std	0.6	0.2	0.4	0.5	0.7	0.9	0.4	0.0	0.7
Subj 14	Avg	16.5	12.3	10.9	9.4	7.1	6.7	3.8	3.1	2.5
	Std	0.6	0.5	1.2	0.6	0.1	0.0	0.2	1.1	0.2
Subj 15	Avg	13.5	10.7	8.3	7.1	7.4	4.5	3.5	2.7	
	Std	0.2	0.1	0.5	0.1	0.0	0.1	0.0	0.1	
Subj 16	Avg	14.9	12.0	10.2	9.1	7.7	5.6	4.7	4.3	2.6
	Std	0.2	0.1	0.4	0.5	0.1	0.2	0.1	0.1	0.1

Table D.2: Mean values (Avg) of the TBOAE latency estimates for each subject at each frequencies. The standard deviation (Std) is also presented. When OAEs could not be detected, the fields were left blank. The results are discussed in section 3.3.5 and plotted on figure 3.9 page 37. Despite a good repeatability for each subject, the OAE latency estimate presents a sizable variation across subjects.

## Effect of tone burst rise time on wave-V latency

---

The tone bursts used to measure OAEs and ABRs were inspired by the experiment of Neely et al. (1988), who used different rise times across frequency in order to have the same amount of BM excitation for each stimulus. Recently, Ruggero and Temchin (2007) claimed that the difference in stimulus rise time in Neely et al. (1988) added a supplementary delay to  $\tau_{wave\ V}$ . They argued that identical tone-burst rise times are necessary to have the synchronous neural firing occurring at the same time for all the stimuli. They base their critics on a study by Heil and Irvine (1997) who studied the delay of the first spike in the auditory nerve. They are correct in their concerns when exciting near threshold levels. The level of the present experiment is, however, much higher (66 dB pe SPL) and the effect of the stimulus rise time is limited. The same Heil & Irvine studies (figure 2E) shows that for tones levels above 60 dB SPL, the first spike latency difference is less than 0.5 ms when the ramp decreases from 4.2 to 1.7 ms. Despite this approximation, the point raised by Ruggero and Temchin (2007) should be addressed. How much of the wave-V latency estimate can be accounted for by the difference in rise time? Is it possible to somehow compensate for this difference between high and low frequencies?

Studies have shown that, for a given tone burst, a decrease of the rise time leads to a decrease of the wave-V latency (Kodera et al., 1977b; Stapells and Picton, 1981; Beattie et al., 1984; Beattie and Torre, 1997). This phenomenon can be explained by a larger spread of the spectrum when the rise time is decreased. This means that the tone bursts will activate a more basal part of the cochlea, inducing earlier nerve firing and

Freq. [kHz]	0.5	0.75	1	1.5	2	3	4	6	8
$\Delta_{waveV}$ [ms] $RT = 5$	0	0.83	1.23	0.94	0.80	0.95	0.99	0.90	0.79
$\Delta_{waveV}$ [ms] $RT = 0.625$	-2.66	-1.59	-0.92	-0.71	-0.60	-0.31	-0.16	-0.05	0

Table E.1: Correcting factor,  $\Delta_{waveV}$ , of the different tone bursts for two different rise times,  $RT = 5$  ms and  $RT = 0.625$  ms. This calculation is based on data published by Stapells and Picton (1981) at 70 dB pe SPL.

leading to earlier wave V (Hecox and Deegan, 1983). If the rise time is increased, this leads to a decrease of wave-V amplitude, due to fewer neurons firing for the narrower stimuli. This is critical since it decreases the detectability and can result in a higher *intra*-subject variability.

However, Stapells and Picton (1981) investigated different aspects of ABRs when using tone bursts as stimuli. Their experiment 5 showed that the wave-V latency,  $\tau_{wave V}$ , increases when the stimulus rise time increases. Based on this experiment, a factor was calculated for  $\tau_{wave V}$  in order to compensate for an equal rise time in two extreme cases: a) all stimuli have a rise time of 5 ms (the original rise time of the 0.5 kHz tone burst) and b) all stimuli have a rise time of 0.625 ms (the 8 kHz tone burst rise time). The correcting factors for each tone burst are presented in table E.1.

To investigate the exact consequence of such correcting factor on the results previously found,  $\hat{\tau}_{BM}^{(ABR)}$  is estimated with the new individual values of  $\tau_{wave V}$ . For each subject  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  are compared in a way similar to that presented in chapter 5, i.e., an ANOVA test was performed to see if twice the new estimate of the BM delay is significantly different from the OAE latency. For ten out of eleven subjects, the p-value is less than 0.05, indicating a significant difference between these two variables. It seems therefore that the “normalisation” of the rise time increases the effect previously observed. It can be concluded that the observation made while comparing  $\tau_{OAE}$  and  $2\hat{\tau}_{BM}^{(ABR)}$  is still valid with an equal rise time for all tone bursts. However, due to the lack of frequency specificity and the possible decrease of the wave V amplitude, it is not recommended to use an equal rise time for all the tone bursts. The bias introduced by choosing equal rise times becomes obvious when looking at the tone bursts  $Q$ -factors in table E.2. This table shows that a similar rise time for all stimuli leads to a very different spread in frequency. For instance, the

Freq. [kHz]	0.5	0.75	1	1.5	2	3	4	6	8
$Q$	4.31	4.05	3.59	5.83	7.19	7.98	7.36	7.43	7.3
$Q_5$	4.31	6.46	7.41	11.1	14.8	25.86	29.63	51.72	59.26
$Q_{0.6}$	0.65	0.80	0.89	1.37	1.79	2.68	3.65	5.48	7.3

Table E.2: Change in the tone-burst  $Q$ -factor between the rise time used in this study and the extreme cases  $RT = 5$  ms and  $RT = 0.625$  ms. This clearly shows that taking a constant rise time across frequency leads to a great variation of the tone bursts  $Q$ -factor.

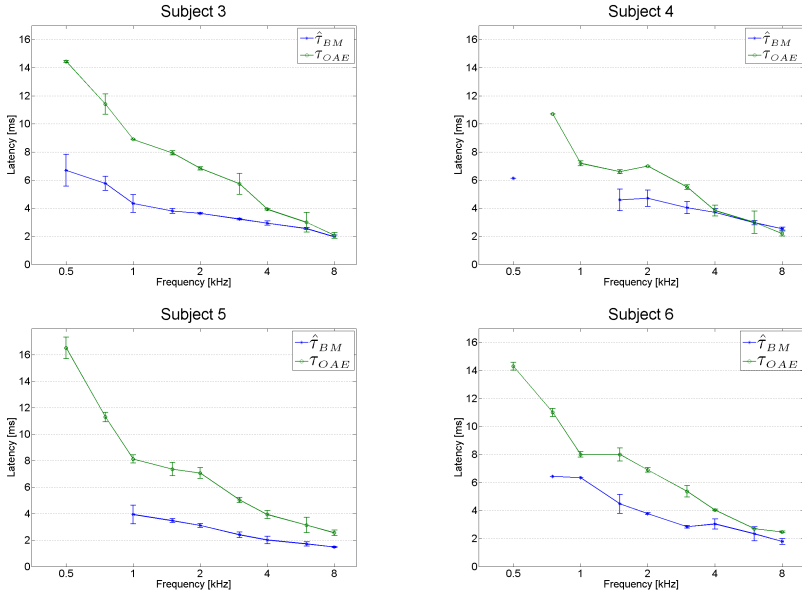
8 kHz tone burst with a 5 ms rise time is too spread in frequency and it will excite a broad area of the BM. On the contrary, the 0.5 kHz TB with 0.625 ms rise time is too narrow and will not lead to a sufficient BM excitation to elicit a nerve response. This demonstrates that the compromise made in the tone burst duration leads to an almost constant  $Q$ -factor which will lead to a similar number of neurons excited for each stimulus.

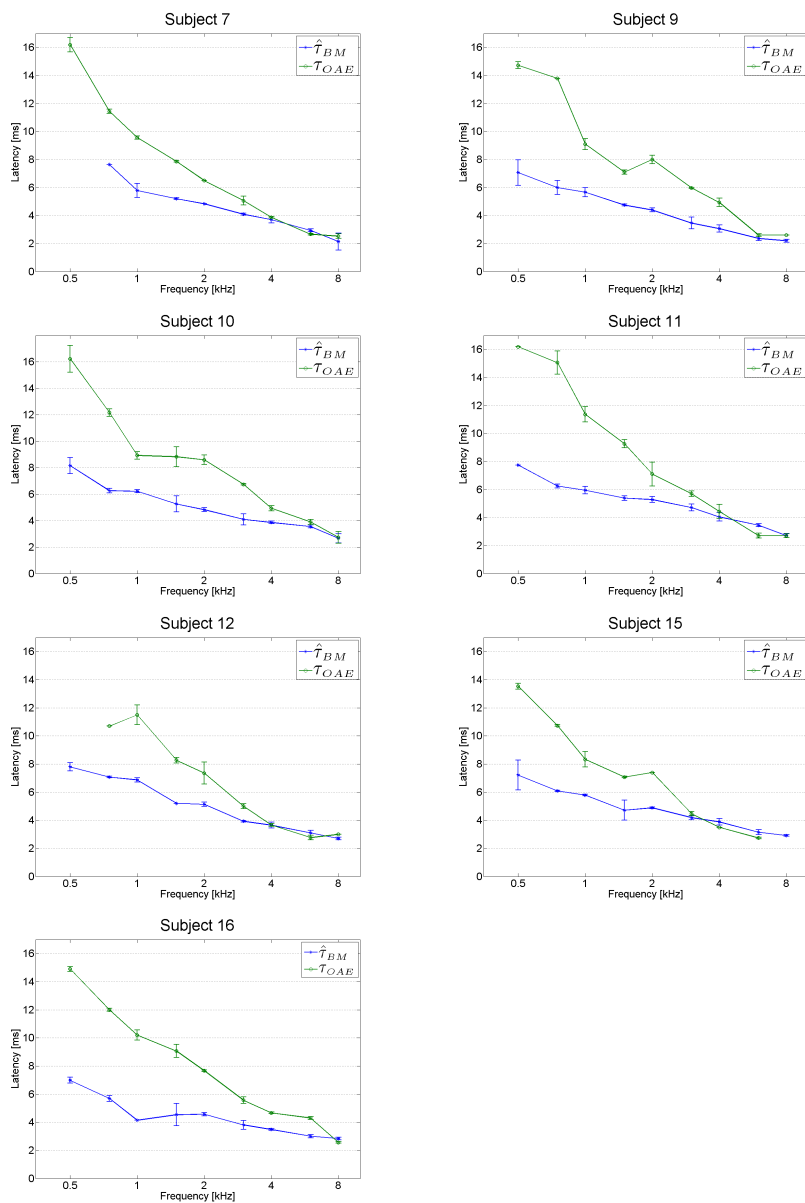


# Acoustic wave theory results

---

In this appendix, the latencies of OAEs ( $\tau_{OAE}$ ) are compared with the latency of the forward travelling wave  $\hat{\tau}_{BM}^{(ABR)}$  estimated from the wave-V latency. This results are discussed in section 5.1.1 page 66.



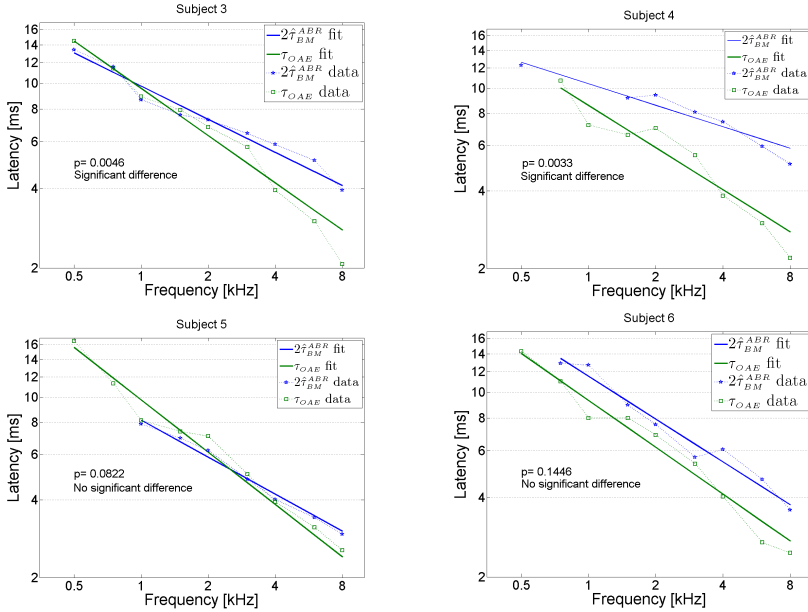


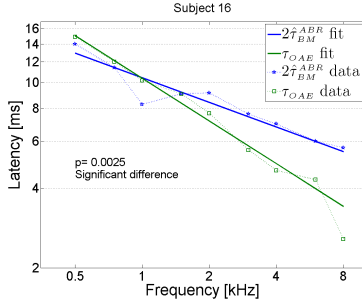
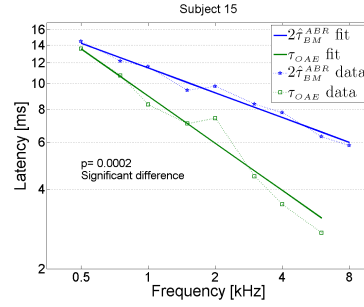
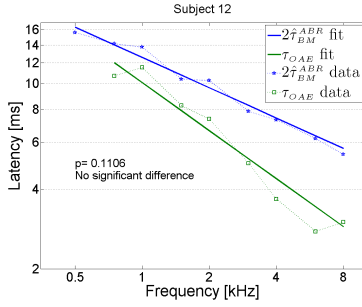
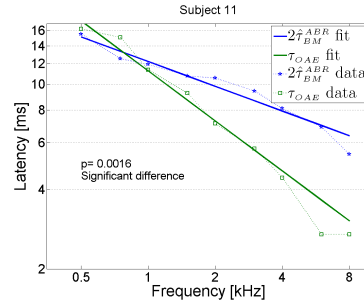
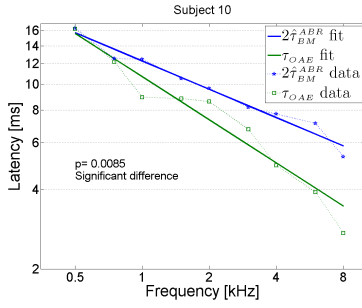
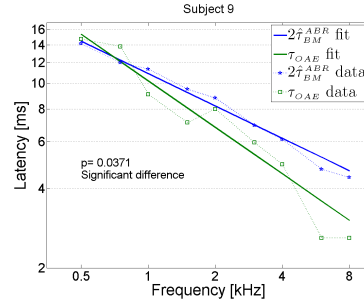
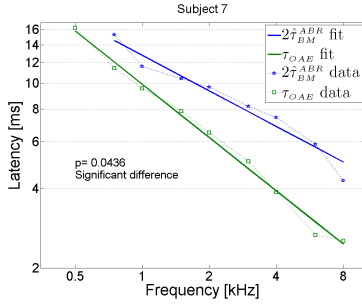


## CRF results

---

This appendix presents the confrontation of the results with the coherent reflection filtering theory. This theory predicts  $\tau_{OAE} = 2\tau_{BM} \cdot \tau_{BM}$  is here estimated from the ABR wave-V latency. Curves were fitted to the data points and an ANOVA test was run to see whether these two estimates  $\tau_{OAE}$  and  $\hat{\tau}_{BM}^{(ABR)}$  are significantly different or not, see section 5.1.2 page 66.

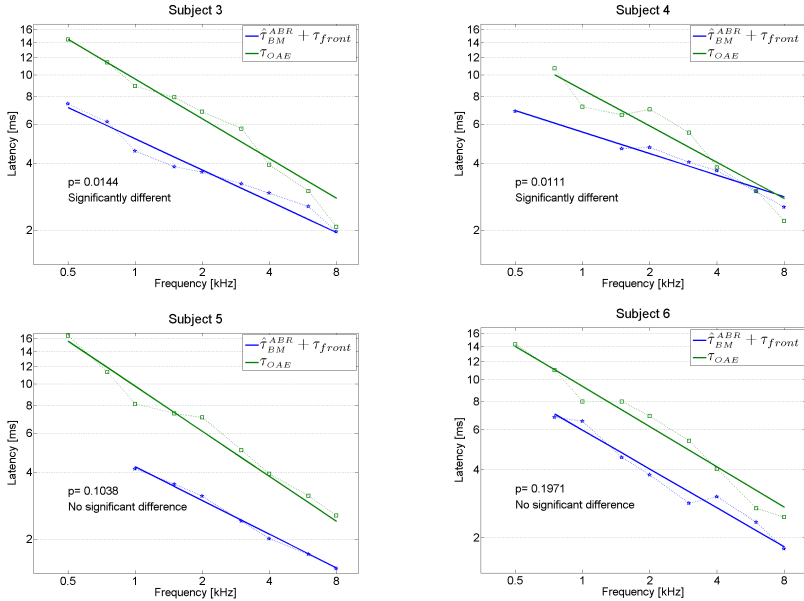


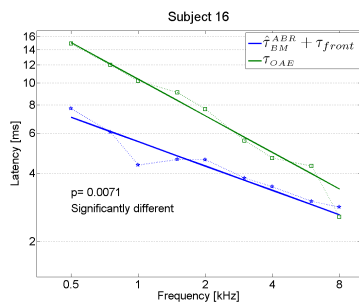
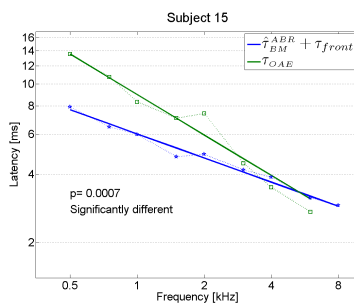
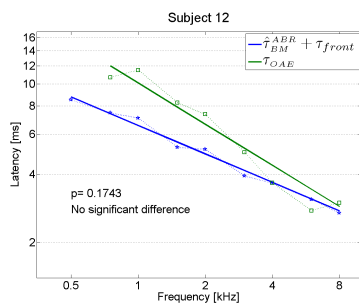
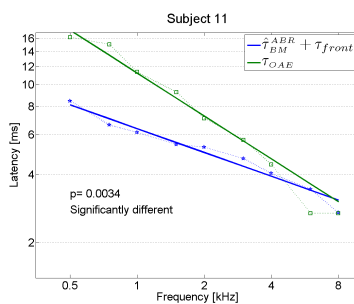
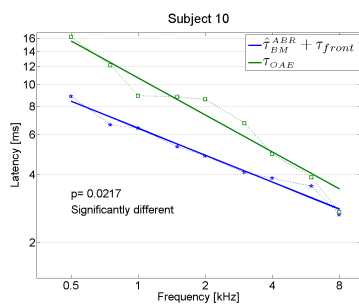
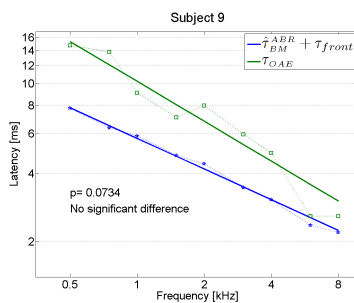
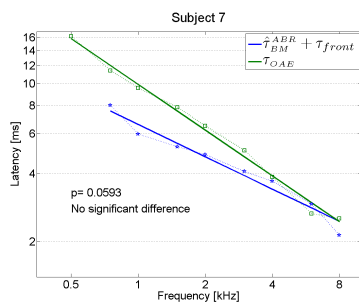


## Front delay results

---

This appendix presents the comparison between the OAE latency  $\tau_{OAE}$  and the estimate given by the signal-front theory, i.e.  $\tau_{front} + \hat{\tau}_{BM}^{(ABR)}$ . As for the CRF theory, an ANOVA test was done to see if the two estimates were significantly different or not. The plot shows the value of  $p$ . More details are given in section 5.1.3 page 67.





## Individual ASSR latencies

---

Frequencies [kHz]			0.5	0.8	1	1.6	4	6.3
Subj 1	M	Avg Std	22.0 0.8	20.5	19.9 0.3	17.9 0.1	17.2 0.1	
Subj 2	F	Avg Std	20.4	22.7	18.8 0.6	17.8 0.03	16.8 0.2	
Subj 3	M	Avg Std	20.2 0.05		18.6 0.1	18.2 0.4	16.7 0.2	16.6 0.2
Subj 4	F	Avg Std	18.7 0.6	16.9 0.5	16.8 0.4	14.8 0.7	14.4 0.1	16.7
Subj 5	F	Avg Std	19.7 0.1	20.3 0.4	17.9 0.4	16.9 0.3	14.5	16.4 1
Subj 6	M		20.8	21.3	21.1	18.6	14.0	
Subj 7	F		20.9	20.4	18.7	18.4		17.6
Subj 8	F	Avg Std	20.2 0.02	20.4	19.3 0.2	17.1 0.1	16.4 0.5	
Subj 9	M	Avg Std	20.8 0.1	20.7	19.7 0.1	18.7 0.1	17.9 0.6	15.9
Subj 10	F		22.0	20.3	17.9	17.3		16.6
Subj 11	F				17.9	15.8	15.7	
Subj 12	M		20.6	20.2	18.4	16.9	16.1	16.3
Subj 13	M	Avg Std	22.5 0.2	21.9 0.1	19.8 0.1	17.8 0.0	16.9 0.2	

Table I.1: Individual ASSR latencies for the 13 subjects. Subjects 1-5,8,9 and 13 were tested twice to investigate the intra-subject variability, the average values and the std are indicated. When the recording did not pass the MSC test, no latency value was calculated, explaining the blanks in the table. In the same way, when data were available only for one run the std is not calculated. M/F indicates if the subject is a male or a female.