

**Auditory brainstem and  
middle-latency responses with  
optimized stimuli:  
Experiments and models**

Vom Fachbereich Physik der Universität Oldenburg  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften (Dr. rer. nat.)  
angenommene Dissertation.

**Oliver Fobel (geb. Wegner)**

geb. am 20.10.1971

in Oldenburg

Erstreferent: Prof. Dr. Dr. Birger Kollmeier

Korreferent: Prof. Dr. Volker Mellert

Tag der Disputation: 23. Januar 2003

# Contents

<b>1</b>	<b>General introduction</b>	<b>7</b>
<b>2</b>	<b>Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion</b>	<b>11</b>
2.1	Introduction . . . . .	12
2.2	The chirp stimulus . . . . .	15
2.3	Method . . . . .	19
2.3.1	Subjects . . . . .	19
2.3.2	Apparatus . . . . .	19
2.3.3	Stimuli and procedure . . . . .	20
2.4	Results . . . . .	21
2.4.1	Click- versus chirp-evoked potentials . . . . .	21
2.4.2	Exact versus approximated chirp . . . . .	23
2.4.3	Stimulation with ramped chirps . . . . .	25
2.4.4	Effects of direction of frequency sweeping . . . . .	26
2.4.5	Effects of spectral composition . . . . .	29
2.5	Discussion . . . . .	30
2.5.1	Stimulus presentation: SL versus peSPL . . . . .	30
2.5.2	Role of wave V behavior – spectral integration . . . . .	31
2.5.3	Assumption of linearity of BM characteristic . . . . .	33
2.6	Summary and conclusions . . . . .	33
<b>3</b>	<b>Frequency specificity of chirp-evoked auditory brainstem responses</b>	<b>37</b>
3.1	Introduction . . . . .	38
3.2	Method . . . . .	41

3.2.1	Subjects . . . . .	41
3.2.2	Apparatus . . . . .	41
3.2.3	Stimuli and procedure . . . . .	42
3.2.4	Experimental masking paradigms . . . . .	44
3.2.5	Statistical analysis . . . . .	45
3.3	Results . . . . .	46
3.3.1	Click- versus chirp-evoked responses using noise masking . . . . .	46
3.3.2	Tone-pulse versus low-frequency chirp-evoked responses . . . . .	50
3.4	Discussion . . . . .	53
3.4.1	Broadband chirp versus click . . . . .	53
3.4.2	Low-frequency chirp versus tone pulse . . . . .	57
3.4.3	Limitations of the chirp . . . . .	59
3.4.4	Relation to other methods of frequency specificity . . . . .	60
3.5	Summary and conclusions . . . . .	61
<b>4</b>	<b>Searching for the optimal stimulus eliciting auditory brainstem responses in humans</b>	<b>63</b>
4.1	Introduction . . . . .	64
4.2	The chirp stimuli . . . . .	67
4.2.1	OAE-based chirp stimulus (O-chirp) . . . . .	67
4.2.2	ABR based chirp stimulus (A-chirp) . . . . .	68
4.2.3	Comparison of the different chirp stimuli . . . . .	69
4.3	Method . . . . .	71
4.3.1	Subjects . . . . .	71
4.3.2	Apparatus . . . . .	71
4.3.3	Stimuli and procedure . . . . .	72
4.3.4	Statistical analysis . . . . .	74
4.4	Results . . . . .	75
4.5	Discussion . . . . .	78
4.6	Summary and conclusions . . . . .	80
<b>5</b>	<b>Modeling auditory evoked middle latency responses (MLR)</b>	<b>81</b>
5.1	Introduction . . . . .	82

---

5.2	Method . . . . .	84
5.2.1	Subjects and stimulation paradigm . . . . .	84
5.2.2	Recording . . . . .	85
5.2.3	Data analysis . . . . .	85
5.3	The model for MLR generation . . . . .	86
5.3.1	The general modeling approach . . . . .	86
5.3.2	The auditory-nerve model . . . . .	87
5.3.3	The unitary response function . . . . .	89
5.4	Results . . . . .	90
5.4.1	Click- and chirp-evoked MLR as a function of level . . . . .	90
5.4.2	MLRs as a function of the click rate . . . . .	92
5.5	Discussion . . . . .	94
5.6	Summary and conclusions . . . . .	98
<b>6</b>	<b>Summary and conclusions</b>	<b>101</b>
<b>A</b>	<b>Experiments on the correlation between psychophysical loudness and auditory brainstem responses</b>	<b>105</b>
A.1	Introduction . . . . .	105
A.2	Methods . . . . .	106
A.2.1	Psychophysics . . . . .	106
A.2.2	Evoked potentials . . . . .	107
A.3	Results . . . . .	107
A.3.1	Psychoacoustical experiments . . . . .	107
A.3.2	Evoked potentials . . . . .	108
A.4	Discussion . . . . .	109
A.5	Summary . . . . .	112
<b>B</b>	<b>On the relationship between auditory evoked potentials and psychophysical loudness</b>	<b>113</b>
B.1	Introduction . . . . .	113
B.2	Methods . . . . .	114
B.3	Results . . . . .	114

---

B.3.1	Experiment 1 . . . . .	114
B.3.2	Experiment 2 . . . . .	115
B.3.3	Experiment 3 . . . . .	117
B.4	Summary and conclusion . . . . .	117
<b>References</b>		<b>119</b>

# Chapter 1

## General introduction

Evoked potentials represent a very powerful noninvasive neurophysiological technique for evaluating the integrity of human sensory processing systems. Stimulus-evoked and event-related neuroelectric recording techniques provide the temporal resolution required for identification of the neural correlates of the various stages of information processing. Auditory evoked potentials (AEP) are used for clinical assessment of both peripheral and central auditory mechanisms. When recorded from the scalp, AEP represent the contribution of neural events that arise from many discrete and neural generating sites along the auditory pathway from the cochlea to the cerebral cortex, which themselves consist of multiple generators. Thus, AEP offer an objective tool to investigate the function of the auditory system. They have been classified according to their neural generators and/or their corresponding latency: auditory brainstem responses (ABR), middle latency responses (MLR) and cortical auditory evoked responses (CAEP). Since the short-latency ABR do not depend on the arousal and vigilance state of the subjects, they are accessible with less effort than the AEP with longer latencies. Therefore, ABR are used in clinical diagnostics for the objective assessment of hearing disorders, which cannot be assessed reliably by other audiological procedures. In addition to these more applied aspects, ABR are also very interesting for the understanding of basic questions related to the signal processing in the auditory periphery, due to their generation sites in the early stages of the auditory pathway. In particular, ABR allow conclusions about the role of the cochlea for the formation of the potential patterns, and, in turn, may serve as a critical test for existing models of cochlear signal processing in human listeners.

It has generally been assumed that the early components of the AEP, like the ABR, are best evoked by stimulation with clicks. A click has a broad power spectrum capable of activating a large portion of the basilar membrane (BM). However, when a transient stimulus progresses apically along the BM, single-unit activity is less synchronous with the preceding activity from basal units (Tsuchitani, 1983) because of temporal delays imposed by the traveling wave. This results in an asynchronous pattern of neural firing along the length of the cochlear partition. In addition, it is likely that activity generated from single units in more synchronous basal regions would be out of phase with activity from some apical units. As a consequence, the combination of phase cancellation and loss of synchronization bias the evoked potential to reflect activity from more basal, high-frequency regions of the cochlea (e.g., Neely *et al.*, 1988). Thus, it can be expected that the click is not the optimal stimulus to be used in recording ABR.

The main goal of the present thesis is to obtain a deeper understanding of the role of cochlear processing for the generation and formation of ABR (and MLR). Several aspects have been in the focus of research in this thesis: The general idea is to develop and evaluate an optimal stimulus for eliciting ABR based on current knowledge about the signal processing in the human cochlea. “Optimal” is meant in terms of producing a higher neural synchronization than can be achieved with other stimuli, including the click. Three different approaches are considered, each of which serves as a basis for the generation of such an optimized stimulus. By comparing the results for the different strategies, the study attempts to work out what aspects of cochlear and retro-cochlear processing are essential for effective generation of a large far-field response. The results are also compared with quantitative predictions using a functional model of the auditory periphery.

As the click, the desired (“optimal”) stimulus must have a wideband frequency spectrum to excite a maximal number of nerve fibers. Since there is temporal dispersion of displacement maxima along the cochlear partition, the temporal spacing of frequency components of the wideband stimulus must be adjusted to provide maximum synchrony of discharges across frequency. A low-frequency tone requires more time to reach its place of maximum displacement, near the apex of the cochlea, than does a high-frequency tone that elicits a maximum closer to the base. Thus, in the hypothesized “ideal” stimulus, the high-frequency components must be delayed relative to the low-frequency components. The acoustic signal must therefore be a rising frequency chirp.



---

In chapter 2, a linear basilar membrane model (de Boer, 1980) is used to calculate the time course of frequency change for a chirp that theoretically produces simultaneous maxima by compensating for travel-time differences along the cochlear partition. This chirp is used to evoke ABR at different stimulation levels. Chirp-evoked responses are directly compared with responses obtained with the “classical” click stimulus and those obtained with a temporally reversed (falling frequency) chirp.

The usefulness of the rising chirp from chapter 2 for retrieving frequency-specific information is examined in chapter 3. This is of particular interest for clinical applications since such frequency-specific responses may serve as a reliable estimate and neural correlate of frequency-specific hearing. In the first experimental series, chirp and click-evoked ABR responses are obtained in the presence of high-pass and notched-noise maskers, for a set of cut-off frequencies of the noise ranging between 0.5 and 8 kHz. In another series of experiments, responses obtained with a low-frequency chirp are compared with those obtained for a tone pulse having comparable duration and magnitude spectrum.

The model-based design for the generation of the rising chirp as suggested in chapter 2 represents only one possible paradigm. In chapter 4, two other strategies for chirp generation are developed. The first one is based on stimulus-frequency otoacoustic emissions (SFOAE) in humans, obtained at a stimulation level of 40 dB SPL. Shera and Guinan (2000) calculated emission group delays from their SFOAE data and related them to BM group delays as a function of characteristic frequency. Based on these estimates of BM group delays, a corresponding chirp stimulus is calculated. In addition, a second approach is tested, which is based on tone-pulse ABR data from Gorga *et al.* (1988). They measured ABR latencies for a wide range of frequencies and levels. Neely *et al.* (1988) used their data to give a prediction of BM group delay in the human cochlea. Their equation, which is a function of level and frequency, is used to develop a level dependent optimized chirp stimulus. The different approaches are compared and corresponding ABR experiments are presented.

In chapter 5, a model for the generation of middle-latency responses (MLR) is introduced. It represents an extension of the model for ABR generation by Dau (2003). Within the model it is assumed that evoked potentials recorded at remote electrodes can be described as convolution of an effective elementary unit waveform (unitary response) with a model-derived stimulus-dependent neural excitation produced by the peripheral auditory-nerve activity. All nonlinearity in the model is located in the peripheral filtering stage while

later processing is linear. Experiments using clicks and chirp stimuli at different stimulation levels and also clicks at different repetition rates are presented. The data are compared with corresponding model predictions to show the capabilities and limitations of the above modeling approach. Particularly, it is investigated to what extent the typically found strong response at a repetition rate of 40 Hz can already be described quantitatively in terms of a linear-system's analysis, or whether additional processes need to be considered.

Finally, the relationship between AEP (particularly the response amplitude) obtained with rising chirps and the corresponding psychophysical sensation of loudness, particularly the loudness function and loudness summation, are investigated. The first results of these experiments are presented and discussed in the appendix of this thesis.

## Chapter 2

# Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion<sup>1</sup>

---

<sup>1</sup> This chapter was published as a paper with the same title written together with Torsten Dau, Volker Mellert and Birger Kollmeier, see [Dau \*et al.\* \(2000\)](#).

## Abstract

This study examines auditory brainstem responses (ABR) elicited by rising frequency chirps. The time course of frequency change for the chirp theoretically produces simultaneous displacement maxima by compensating for travel-time differences along the cochlear partition. This broadband chirp was derived on the basis of a linear cochlea model [de Boer, “Auditory physics. Physical principles in hearing theory I,” *Phys. Rep.* **62**, 87–174 (1980)]. Responses elicited by the broadband chirp show a larger wave-V amplitude than do click-evoked responses for most stimulation levels tested. This result is in contrast to the general hypothesis that the ABR is an electrophysiological event most effectively evoked by the onset or offset of an acoustic stimulus, and unaffected by further stimulation. The use of this rising frequency chirp enables the inclusion of activity from lower frequency regions, whereas with a click, synchrony is decreased in accordance with decreasing traveling velocity in the apical region. The use of a temporally reversed (falling) chirp leads to a further decrease in synchrony as reflected in ABR responses that are smaller than those from a click. These results are compatible with earlier experimental results from recordings of compound action potentials (CAP) [Shore and Nuttall, “High synchrony compound action potentials evoked by rising frequency-swept tonebursts,” *J. Acoust. Soc. Am.* **78**(4), 1286–1295 (1985)] reflecting activity at the level of the auditory nerve. Since the ABR components considered here presumably reflect neural response from the brainstem, the effect of an optimized synchronization at the peripheral level can also be observed at the brainstem level. The rising chirp may therefore be of clinical use in assessing the integrity of the entire peripheral organ and not just its basal end.

## 2.1 Introduction

It is generally assumed that the conventional auditory brainstem response (ABR) is an electrophysiological event evoked by the *onset* of an acoustic stimulus. Whether the stimulus is an acoustic click, tone pip, tone burst, or noise burst, the ABR is assumed to be effectively evoked by the first few milliseconds of the stimulus, and is generally unaffected by further stimulation (e.g., [Hecox \*et al.\*, 1976](#); [Kodera \*et al.\*, 1977](#); [Debruyne and Forrez, 1982](#); [Gorga and Thornton, 1989](#); [van Campen \*et al.\*, 1997](#)). Because of its abrupt onset, the acoustic click

is often thought to be an ideal stimulus for eliciting a detectable ABR. Clicks or impulsive stimuli are also used under the assumption that their wide spectral spread, inherent in transient signals, elicits synchronous discharges from a large proportion of cochlear fibers (e.g., Kodera *et al.*, 1977; Gorga and Thornton, 1989; van der Drift *et al.*, 1988a,b).

Additionally, it is generally observed that if a long-duration tone burst ( $> 8$  ms) is employed, a second response can be evoked at stimulus *offset*. This offset response resembles onset ABR morphology and occurs within 8 ms after stimulus offset. It has been termed “offset ABR” or “off potential of the brainstem” (Kodera *et al.*, 1977; Brinkmann and Scherg, 1979).

When a transient stimulus progresses apically along the basilar membrane (BM), single-unit activity is less synchronous with the preceding activity from basal units (Tsuchitani, 1983) because of temporal delays imposed by the traveling wave. This results in an asynchronous pattern of auditory-nerve-fiber firing along the length of the cochlear partition. In addition, it is likely that activity generated from the single units in more synchronous basal regions would be out of phase with activity from some apical units. As a consequence, the combination of phase cancellation and loss of synchronization *bias* the evoked potential to reflect activity from more basal, high-frequency regions of the cochlea (e.g., Neely *et al.*, 1988).

More evidence about the interaction between basilar-membrane dispersion and the synchrony of neural responses can be derived from studies on the compound action potential (CAP) which represents auditory-nerve activity. When stimulated with a click, only auditory-nerve units tuned above 2–3 kHz contribute to synchronous activity in the  $N_1P_1$  complex (Dolan *et al.*, 1983; Evans and Elberling, 1982). In order to determine if cochlear units tuned below 2–3 kHz could be recruited into the CAP response, Shore and Nuttall (1985) used tone bursts of exponentially rising frequency to hypothetically activate synchronous discharges of VIIIth-nerve fibers along the length of the cochlear partition. Their equations defining the frequency chirps were calculated to be the inverse of the delay-line characteristic of the guinea pig partition. Shore and Nuttall (1985) recorded CAPs in response to the rising chirp and compared them to CAP waveforms evoked by corresponding falling chirps as well as clicks. Their analysis of the CAP waveforms showed narrower  $N_1$  widths and larger  $N_1$  and  $P_1$  amplitudes for rising sweeps when compared to falling sweeps. Their results supported the hypothesis underlying the derivation of the rising sweep: spectral

energy with the appropriate temporal organization, determined by basilar membrane traveling wave properties, increases CAP synchrony. In a later study, Shore *et al.* (1987) provided evidence that the timing of discharges in the ventral cochlear nuclei (VCN) reflects cochlear partition motion as demonstrated for VIIIth-nerve fibers and inner hair cells (Brugge *et al.*, 1969; Geisler *et al.*, 1974; Rose *et al.*, 1971; Russell and Sellick, 1978; Sellick *et al.*, 1982). However, unlike VIIIth-nerve fibers, responses of VCN neurons to rapid frequency sweeps were more complex, showing directional preferences.

The present paper followed the same strategy of generating an “optimized” stimulus causing maximal synchronous activation at the level of VIIIth nerve, but deals with brainstem recordings in human subjects. The latencies of the brainstem potentials can be separated into mechanical and neural components. The mechanical component is due to mechanical BM travel time, and varies with intensity and frequency in an orderly manner, while the remaining neural component is assumed to be independent of both intensity and frequency (e.g., Neely *et al.*, 1988). We attempted to compensate for the frequency-dependent mechanical component in order to increase synchrony at a peripheral level, which may also lead to increased synchrony at higher stations in the brainstem. Our question was whether such a stimulus would be appropriate and effective for ABR recordings. Of course, there is a large difference between events in single-unit electrical fields, and the signals which are recorded by electrodes which are remote from the neural sources. Single-unit electrical fields are rapidly attenuated in the extracellular space and are unmeasurable at more than a few millimeters distance. Also, the effectiveness of neural centers as dipole generators producing a detectable far-field response depends on the number of involved neural sources and on morphological features such as dendritic orientation. However, the mechanical component of BM travel time should affect single-unit-electrical-field responses and whole-ensemble-far-field responses in a similar way. Hence, the time-frequency distribution of a stimulus can be expected to have a distinct effect on ABR.

ABRs elicited by broadband, frequency-sweeping stimuli are compared with click-evoked responses. The underlying chirp stimulus was generated on the basis of the (linear) basilar-membrane model by de Boer (1980).

## 2.2 The chirp stimulus

The equations describing the stimulus were derived based on the following considerations: (i) since the mechanical properties of the cochlear partition result in a spatial separation of frequency components of an acoustic signal, the desired stimulus must have a wideband frequency spectrum to excite a maximal number of nerve fibers (see also [Shore and Nuttall, 1985](#)). (ii) Since there is also a *temporal* dispersion of displacement maxima along the cochlear partition, the temporal spacing of frequency components of the wideband signal must be adjusted to provide maximum synchrony of discharge across frequency. A low-frequency tone requires more time to reach its place of maximum displacement, near the apex of the cochlea, than does a higher-frequency tone that elicits a maximum closer to the base. The idea is to generate a stimulus in which the high-frequency components are delayed relative to the low-frequency components by an appropriate amount. This should produce *synchronous* displacement maxima and neural discharges resulting from *all* frequency components. The acoustic signal must therefore be a rising frequency chirp. The time course of the chirp developed in the present study is determined by the traveling wave velocity along the partition as derived by de Boer (1980), and the functional relationship between stimulus frequency and place of maximum displacement ([Greenwood, 1990](#)).

[De Boer \(1980\)](#) developed a cochlear model in which – as physical simplifications – he assumed that the fluids of the canals around the basilar membrane would be incompressible and that all viscosity effects were negligible. All movements were assumed to be so small that the fluid as well as the basilar membrane (BM) operate linearly. All time-dependent variables were considered to vary as  $e^{i\omega t}$ , with  $\omega$  representing radian frequency. Since the dynamics of the BM are certainly nonlinear, in some conditions this linear approach must be considered as a first-order approximation. It was further assumed by [de Boer \(1980\)](#) that various parts of the BM are not mechanically coupled to each other and that all coupling occurs via the surrounding fluid. De Boer described the mechanics of the cochlear partition by a single function of the coordinate  $x$ , the impedance  $\xi(x)$ , which is dominated by a stiffness term  $c(x)/i\omega$ . The fluid movements in the other two directions were assumed not to contribute to the mechanical pattern of movement of the cochlear partition.

The wave equation for the hydro-mechanical problem was then given by

$$\frac{d^2}{dx^2} \psi(x) - \frac{2i\omega\rho}{h(x)\xi(x)} \psi(x) = 0, \quad (2.1)$$

with  $\psi(x)$  as the wave function,  $h(x)$  as the “effective” height of the scala, and the density  $\rho$ . The impedance  $\xi$  is the critical factor and is composed of three parts, a mass part, a resistance part, and a stiffness part:

$$\xi(x) = i\omega m(x) + r(x) + \frac{c(x)}{i\omega}. \quad (2.2)$$

The mass term  $m(x)$  does not depend much on  $x$  in the cochlea, but the stiffness  $c(x)$  varies over a large range as a function of  $x$  (e.g., von Békésy, 1960). It is assumed in the following that *only* stiffness contributes to  $\xi(x)$ , while mass and resistance  $r(x)$  are neglected. Due to the variations of the stiffness, the velocity of propagation depends strongly on  $x$ . The stiffness was shown to be mainly responsible for the occurrence of traveling waves propagating along the cochlear partition. De Boer developed the *exponential model* assuming that  $c(x)$  varies approximately as an exponential function of  $x$ :  $c(x) = C_0 e^{-\alpha x}$ . This results in the wave equation:

$$\frac{d^2}{dx^2} \psi(x) + D_0^2 e^{\alpha x} \psi(x) = 0, \quad (2.3)$$

where  $D_0 = \omega \sqrt{2\rho/hC_0}$ . The solution of this equation was given as

$$\psi(x) = \arctan \frac{Y_0(z)}{J_0(z)}, \quad (2.4)$$

where  $z = (2D_0/\alpha)e^{\alpha x/2}$ , and  $J_0$  and  $Y_0$  represent Bessel and Weber function of zero order, respectively. This leads to the local propagation constant

$$k(x) = \frac{d}{dx} \psi(x) = \frac{\frac{\alpha}{\pi}}{J_0^2(z) + Y_0^2(z)}. \quad (2.5)$$

The speed of propagation is generally given by  $\gamma(x) = \omega/k(x)$ . The exact expression for the exponential model therefore is:

$$\gamma(x) = \frac{\pi\omega}{\alpha} (J_0^2(z) + Y_0^2(z)). \quad (2.6)$$

This expression shows dispersion with respect to frequency since  $\gamma$  is dependent on  $\omega$ .

The above equations are used in the present paper to generate the chirp stimulus that compensates dispersion on the BM. The propagation time  $t_\omega(x)$  needed to arrive at the place of resonance  $x$  is given by:

$$t_\omega(x) = \int_0^x \frac{1}{\gamma} dx' = \frac{1}{\omega} \int_0^x k(x') dx' = \frac{1}{\omega} \int_0^x \frac{d}{dx'} \psi(x') dx' = \frac{1}{\omega} (\psi(x) - \psi(0)), \quad (2.7)$$



leading to:

$$t_\omega(x) = \frac{1}{\omega} \left( \arctan \frac{Y_0(z(x))}{J_0(z(x))} - \arctan \frac{Y_0(z(0))}{J_0(z(0))} \right). \quad (2.8)$$

For the frequency-place transformation, the mapping proposed by [Greenwood \(1990\)](#) was used:

$$x = x(f) = L - \tilde{c} \log(af + 1) = L - c \ln(af + 1), \quad (2.9)$$

with  $a = 0.006046 \text{ Hz}^{-1}$ ,  $\tilde{c} = 16.7 \text{ mm}$ ,  $c = \tilde{c}/\ln 10$ , and  $L = 34.85 \text{ mm}$  representing BM length. It follows that

$$z(x) = z(x(f)) = \frac{2D_0}{\alpha} e^{\frac{\alpha}{2}[L-c\ln(af+1)]} = \frac{2D_0}{\alpha} (af + 1)^{-\frac{\alpha}{2}c} e^{\frac{\alpha}{2}L}. \quad (2.10)$$

Thus  $t_\omega(x)$  is given by:

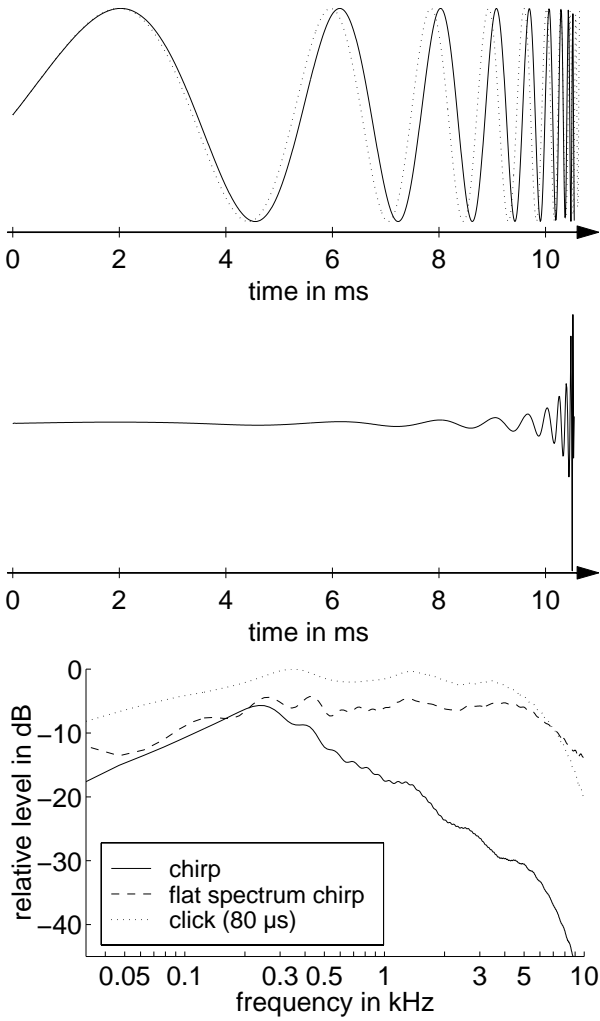
$$t_\omega(x) = t_\omega(x(f)) = \frac{1}{\omega} \left( \arctan \frac{Y_0 \left[ \frac{2D_0}{\alpha} (af + 1)^{-\frac{\alpha}{2}c} e^{\frac{\alpha}{2}L} \right]}{J_0 \left[ \frac{2D_0}{\alpha} (af + 1)^{-\frac{\alpha}{2}c} e^{\frac{\alpha}{2}L} \right]} - \arctan \frac{Y_0 \left( \frac{2D_0}{\alpha} \right)}{J_0 \left( \frac{2D_0}{\alpha} \right)} \right). \quad (2.11)$$

Using the variable transformation  $t \rightarrow t_0 - t$ , and with  $\kappa := (4\pi/\alpha)\sqrt{2\rho/hC_0} \cdot e^{(\alpha/2)L}$ , the function  $t = t(f)$  for the “optimal” input frequency  $\omega = 2\pi f$  is given by:

$$t_0 - t(f) = \frac{1}{2\pi f} \left( \arctan \frac{Y_0 [f\kappa(af + 1)^{-\frac{\alpha}{2}c}]}{J_0 [f\kappa(af + 1)^{-\frac{\alpha}{2}c}]} - \arctan \frac{Y_0 \left( \frac{2D_0}{\alpha} \right)}{J_0 \left( \frac{2D_0}{\alpha} \right)} \right). \quad (2.12)$$

From this relation, the inverse function  $f(t) = t^{-1}(f)$  was derived numerically. This function for the change of the instantaneous frequency was then integrated over time to derive the instantaneous phase  $\varphi = 2\pi \int_0^t f(t') dt'$  of the resulting chirp, which has the general form  $s(t) = A(t) \sin(\varphi(t) - \varphi(t_0))$ . This is referred to as the “exact chirp” throughout this paper (cf. Fig. 2.1, top panel, solid curve).

If one uses only an *asymptotic* expression for the propagation constant, namely  $k(x) \approx D_0 e^{\alpha x/2}$ , the speed of propagation results in  $\gamma(x) \approx \sqrt{(hC_0/2\rho)} \cdot e^{-\alpha x/2}$ , which is independent of  $\omega$  so that there is no dispersion with respect to frequency. The asymptotic expression agrees well with the exact one for frequencies higher than about 5 kHz [for details, see [de Boer \(1980\)](#), p. 147]. For lower frequencies, some small deviations occur for low  $x$ -values, i.e. near the cochlear windows. In this region, the asymptotic expression does not hold true



**Figure 2.1:** Top panel: waveform of the broadband rising (0.1–10.4 kHz) chirp stimulus. The equations defining the chirp were calculated to be the inverse of the delay-line characteristic of the cochlear partition on the basis of the linear cochlea model by *de Boer (1980)*. The solid curve represents the exact chirp, the dotted curve shows the approximated chirp (for details, see text). Middle panel: waveform of a modified chirp referred to as “flat-spectrum chirp” whose phase characteristic is the same as that of the original (exact) chirp. Bottom panel: acoustic spectra of chirp (solid curve), click stimulus (dotted curve), and flat-spectrum chirp (dashed curve), as used in the present study (for details, see text).

any more; however, the extent of this effect is not very large. For the asymptotic case, the instantaneous frequency  $f(t) = t^{-1}(f)$  can be easily derived analytically. With

$$\psi(x) = \int_0^x k(x') dx' + \psi(0) = D_0 \frac{2}{\alpha} \left( e^{\frac{\alpha x}{2}} - 1 \right) + \psi(0), \quad (2.13)$$

and Eqn. 2.7, the travel time  $t(f)$  to the resonance place  $x(f)$  is directly given by:

$$\begin{aligned} t &= \frac{1}{\omega} D_0 \frac{2}{\alpha} \left( e^{\frac{\alpha x}{2}} - 1 \right) \\ &= \frac{1}{\omega} D_0 \frac{2}{\alpha} \left( e^{\frac{\alpha x}{2} (L - c \ln(af+1))} - 1 \right) \\ &= \frac{2}{\alpha} \sqrt{\frac{2\rho}{hC_0}} \left( (af+1)^{-\frac{\alpha c}{2}} e^{\frac{\alpha L}{2}} - 1 \right). \end{aligned} \quad (2.14)$$

With  $t \rightarrow t_0 - t$ , and  $\beta := (2/\alpha)\sqrt{2\rho/hC_0}$ , the function  $f(t)$  is given by:

$$f = \frac{1}{a} \left( \left[ e^{\frac{\alpha}{2}L} \left( 1 + \frac{t_0 - t}{\beta} \right) \right]^{-\frac{2}{\alpha c}} - 1 \right). \quad (2.15)$$

From this, the instantaneous phase  $\varphi$  and the resulting chirp  $s(t)$  can easily be derived as above. This is referred to as the ‘‘approximated chirp’’ in the rest of this paper (cf. Fig. 2.1, top panel, dotted curve).

## 2.3 Method

### 2.3.1 Subjects

Ten normal-hearing subjects (audiometric thresholds 15 dB HL or better) with no history of hearing problems were chosen: two females and eight males. The subjects were between 21 and 35 years of age, and were either paid or volunteered for the experiment.

### 2.3.2 Apparatus

The experiments were carried out with a PC-based computer system which controlled stimulus presentation and recording of evoked potentials. A DSP-card (Ariel DSP32C) converted the digitally generated stimulus (25 kHz, 16 bit) to an analog waveform. The output of the DSP card was connected to a digitally controlled audiometric amplifier, which presented the stimulus through an insert earphone (Etymotic Research ER-2) to the subject.

Electroencephalic activity was recorded from the scalp via silver/silver chloride electrodes, attached to the vertex (positive) and the ipsilateral mastoid (negative). The forehead served as the site for the ground electrode. Interelectrode impedance was maintained below 5 k $\Omega$ . Responses were amplified (80 dB) and bandpass filtered (95–1640 Hz, 6 dB/Oct.) with an commercially available ABR preamplifier (Hortmann Neurootometrie).<sup>2</sup> Extra amplification (Kemo VBF/40) was used to reach the optimum range for the A/D-converter.

<sup>2</sup>In the official data sheet of the preamplifier, a ‘‘hard-wired’’ high-pass cut-off frequency of 30 Hz is given. Unfortunately, we could not replicate this value and found a 3-dB cut-off of 95 Hz. The problem is that this setting will cut out a substantial portion of the wave-V amplitude which results in smaller responses overall, particularly for responses from lower frequency stimulus energy. Since the chirp has much of its energy in the low-frequency region, one can expect that an even larger chirp-evoked wave-V amplitude than observed in the present study will be obtained with a more appropriate filter setting.

This amplification was in the range from 10 to 16 dB, resulting in a total amplification of 90–96 dB. The amplified signal was digitized by the DSP-card (25 kHz, 16 bit), which also performed artifact rejection and signal averaging. Responses were recorded for 26 ms following the stimulus onset.

### 2.3.3 Stimuli and procedure

Broadband chirps as described in Sec. 2.2 were used as stimuli. The chirps started and ended with zero amplitude. If not explicitly stated otherwise, no windowing was applied to the stimuli. Chirp-evoked potentials were compared with click-evoked responses. The click had a duration of 80  $\mu$ s.

The top panel of Fig. 2.1 shows the waveforms of the exact (solid curve) and the approximated chirp (dotted curve), both derived in Sec. 2.2. The stimuli have a flat temporal envelope. Since the value for the speed of propagation  $\gamma$  for lower frequencies is lower for the exact chirp (at places near the cochlear windows), it has a slightly longer duration (10.52 ms instead of 10.48 ms for the approximated chirp). Since for both chirps the instantaneous frequency changes slowly at low frequencies relative to the changes in the high-frequency region, their spectra are dominated by the low frequencies. This is shown in the lower panel of Fig. 2.1 (solid curve). The (acoustic) magnitude spectrum decreases continuously with increasing frequency. The dashed curve in the lower panel indicates the spectrum of a modified chirp which is used later in the study. This modified chirp has a flat amplitude spectrum corresponding to that of the click (dotted curve), while the phase characteristic is the same as that of the original (exact) chirp. The spectra were obtained (at the same sensation level of the stimuli) by coupling the ER-2 insert earphone to a Brüel and Kjær ear simulator (type 4157) with a 1/2-in. condenser microphone (type 4134), a 2669 preamplifier, and a 2610 measuring amplifier. The spectra were obtained from FFTs of 100-trial time-domain averages of the stimulus over an analysis time of 64 ms using a sampling rate of 25 kHz (Stanford Research Systems SR780). The waveforms were not windowed prior to FFT. The middle panel of Fig. 2.1 shows the corresponding temporal waveform of the modified chirp, referred to as the “flat-spectrum chirp” throughout this paper. This stimulus starts with very small amplitudes at low frequencies and increases nonlinearly in amplitude with time.

The subject lay on a couch in an electrically-shielded, sound-proof room, and electrodes were attached. The subject was instructed to keep movement at a minimum, and to sleep

if possible. The lights were turned out at the beginning of the session. Each session lasted between one and two hours, depending on the subject's ability to remain still. The ear of stimulation was chosen randomly, i.e., for each subject one ear was chosen and then maintained. The acoustic signals were delivered at a mean repetition rate of 20 Hz for all stimulus conditions. A temporal jitter of  $\pm 2$  ms was introduced to minimize response superimposition from preceding stimuli. Thus, the resulting interstimulus interval (ISI) was equally distributed between 48 and 52 ms. Each trial consisted of 1000 to 4000 averages, depending on the quality of the response. For each stimulus condition, two independent trials were stored in separate buffers. These are illustrated as superimposed waveforms in the figures to show response replicability.

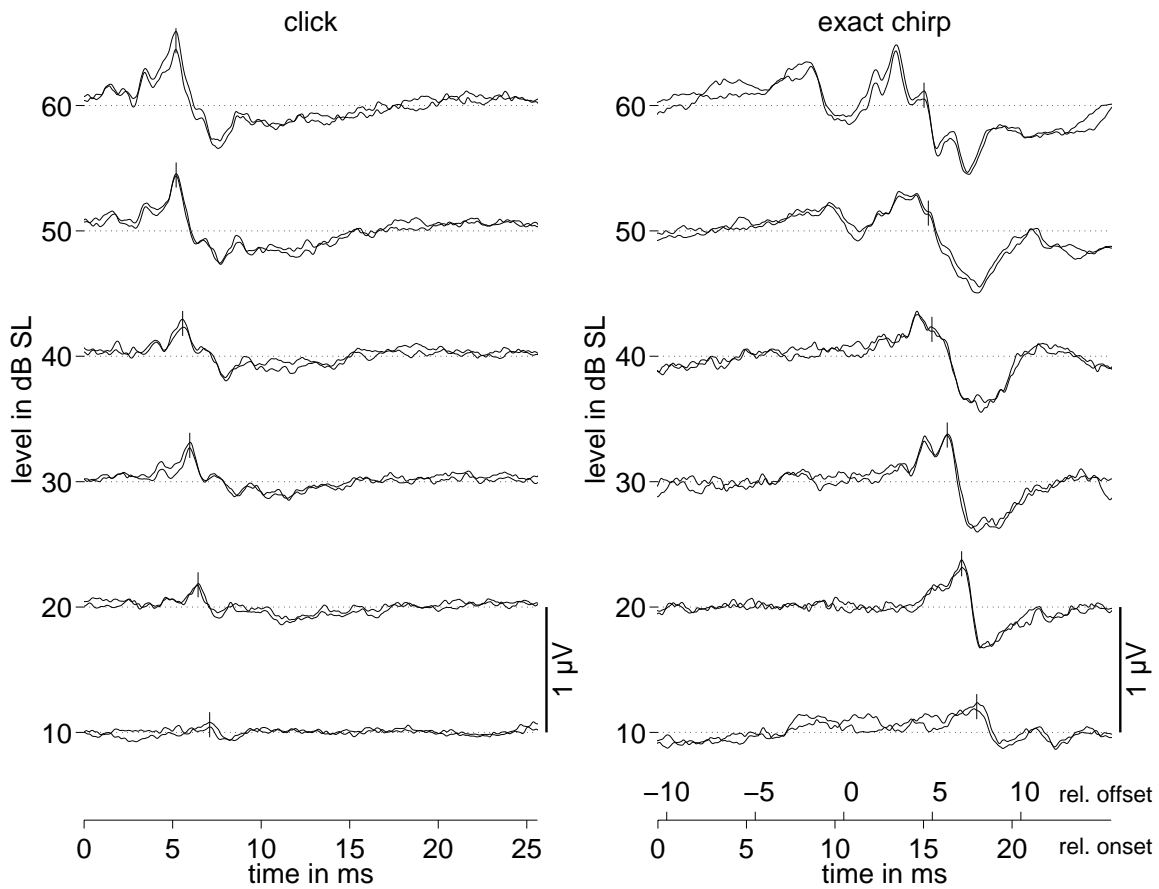
First, to determine the sensation level (SL), both of the click and the chirp stimulus, the absolute hearing thresholds were measured individually with an adaptive 2-AFC procedure. At the beginning of each ABR-recording session, the first trial was a 60 dB SL presentation of a stimulus. Then intensity was decreased in steps of 10 dB down to 10 dB SL. At the same sensation level, chirp and click represent nearly the same root-mean-square (RMS) value (if calculated across the same temporal interval of 10.5 ms).

Wave-V (peak-to-peak) amplitude was analyzed in the different stimulus and level conditions. The amplitude was measured from the peak to the largest negativity following it. For each stimulus and level condition, wave-V amplitude was averaged across subjects. A Wilcoxon matched-pairs signed-rank test ( $\alpha = 0.05$ ) was used to verify whether the response amplitude differed significantly for the two stimuli. Throughout the present paper, responses are plotted for one exemplary subject (CR). Average data for the wave-V amplitude are given in a summary figure (Fig. 2.3).

## 2.4 Results

### 2.4.1 Click- versus chirp-evoked potentials

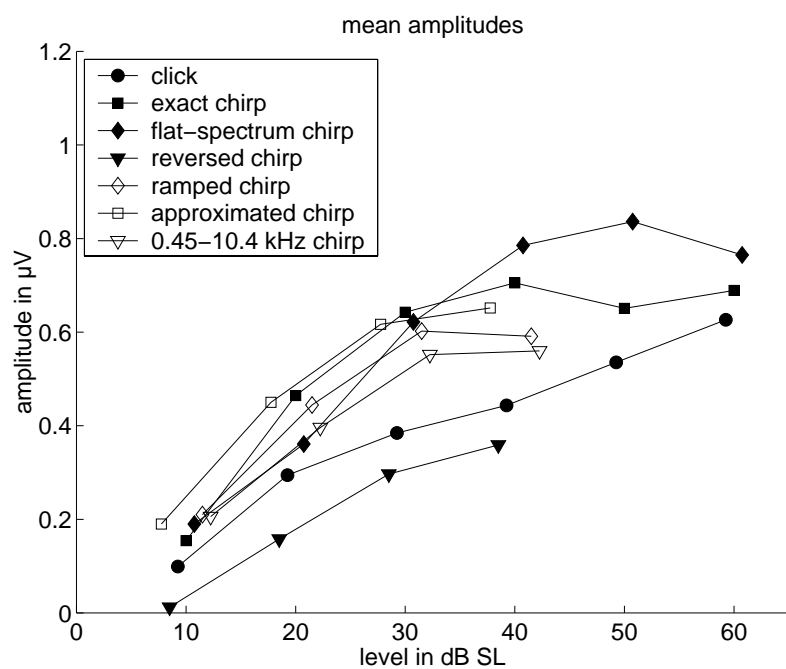
Figure 2.2 shows the ABR for subject CR obtained with a click (left panel) and a rising broadband chirp (right panel), respectively. Responses for different stimulus levels are shown on separate axes displaced along the ordinate and labeled with the sensation level (dB SL). For the click stimulus, the abscissa represents recording time relative to click-onset. In the case



**Figure 2.2:** ABR from subject CR, evoked by a click (left panel) and a broadband chirp (right panel). The stimulation level varied from 10 to 60 dB SL, as indicated. Waveforms are the average of 2000 responses. At each level, two waveforms are superimposed to show response replicability. The small vertical line indicates wave-V peak. Stimulus presentation rate was 20/s.

of the chirp stimulus, a dual abscissa is used representing recording time relative to stimulus onset and offset. Wave-V peak is marked by small vertical bars for both stimuli. It can be seen in the figure that the wave-V latencies for the two stimuli, relative to stimulus onset, are shifted by the duration of the chirp stimulus which equals 10.5 ms. Thus, the latency values relative to stimulus offset are the same in both conditions. The key observation is that the wave-V amplitude is typically larger for chirp stimulation than for click stimulation. For subject CR, the difference is large at stimulation levels of 10–40 dB SL, but is less pronounced at 50 dB SL. At 60 dB SL, for this subject, the click response is slightly larger than the chirp response. At the two highest stimulation levels, earlier activity in response

to the chirp becomes visible with a first response peak at about 8–9 ms after chirp onset. These observations at high levels are probably due to cochlear upward spread of excitation, a well-known phenomenon from many other studies in this field. At high levels, the early low-frequency energy in the chirp stimulates basal regions and produces a response.



**Figure 2.3:** Average ABR data for wave-V amplitude, as a function of the stimulation level. Different symbols indicate different stimulus conditions. ●: click; ■: exact chirp; ◆: flat-spectrum chirp; ▼: reversed chirp; ◇: ramped chirp; □: approximated chirp; ▽: 0.45–10.4 kHz chirp. For better visibility, symbols are slightly shifted along the abscissa.

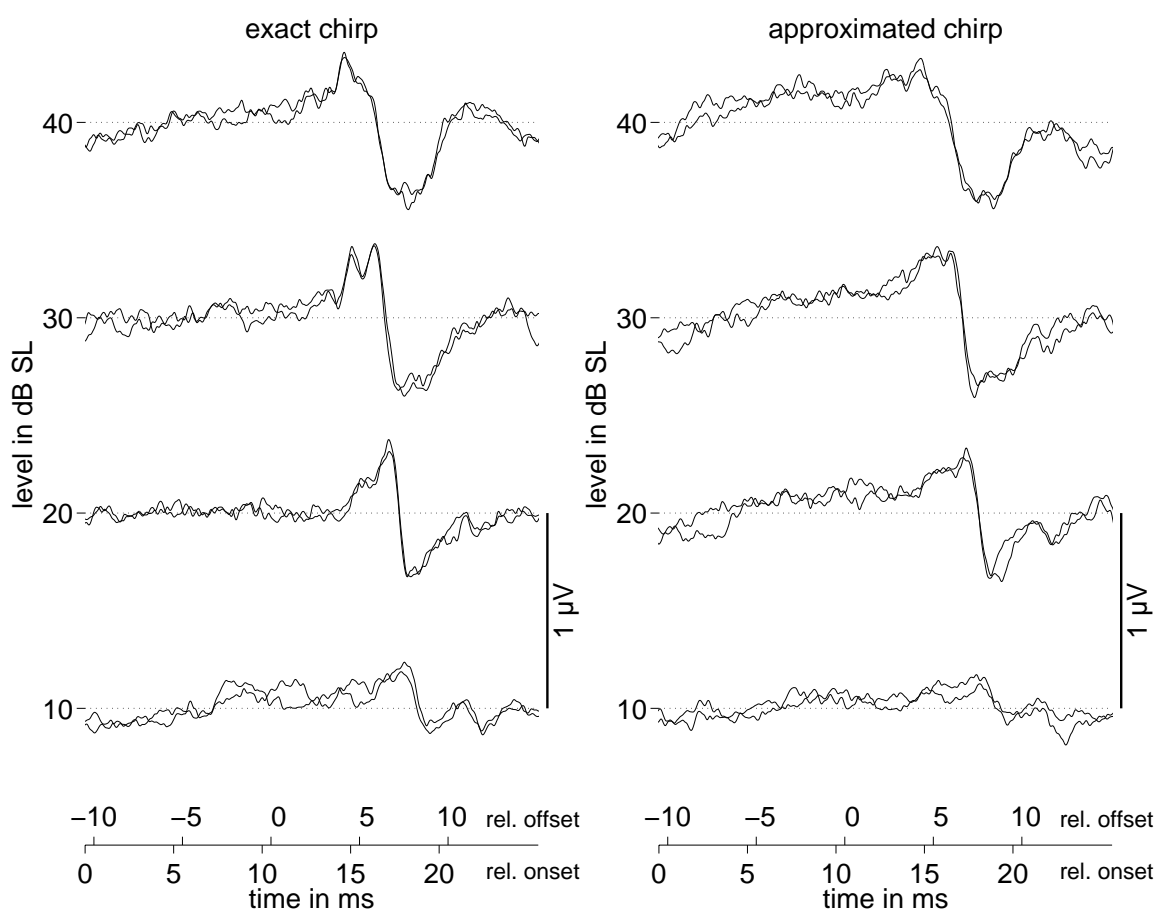
Figure 2.3 shows the wave-V amplitude obtained with different stimuli, including the chirp (filled boxes) and the click (filled circles), averaged across subjects. Amplitude values are plotted as a function of stimulation level. Wave-V amplitude was significantly larger ( $p < 0.05$ ;  $N = 10$ ) for the chirp than for the click, for the levels of 20–40 dB SL. For 50 and 60 dB SL, the average wave-V amplitude was still larger for the chirp than for the click, but the difference was not significant ( $p > 0.05$ ). For the lowest stimulation level, 10 dB SL, four of the subjects showed no clear wave-V peak in either the chirp or in the click condition. The number of the remaining subjects was too small to reveal a significant difference in the ABR.

### 2.4.2 Exact versus approximated chirp

On the basis of the exponential model (de Boer, 1980) reviewed in Sec. 2.2 the generation of a stimulus compensating BM dispersion were derived (see Fig. 2.1). The question is whether

the difference in the time course of the two stimuli is of relevance for the corresponding evoked brainstem potentials.

Figure 2.4 shows ABR for subject CR elicited by the exact chirp (left panel) and by the approximated chirp (right panel). The stimulation level ranged from 10 to 40 dB SL in each case. The potentials are almost identical in both conditions. The average data across subjects for the approximated chirp are indicated as open boxes in Fig. 2.3. Wave-V amplitude does not differ significantly between the two stimuli for all levels ( $p > 0.05$ ;  $N = 6$ ).



**Figure 2.4:** ABR from subject CR, elicited by the exact broadband chirp (left panel) and the approximated chirp (right panel). Parameters as in Fig. 2.2, but only for stimulation levels of 10–40 dB SL.



### 2.4.3 Stimulation with ramped chirps

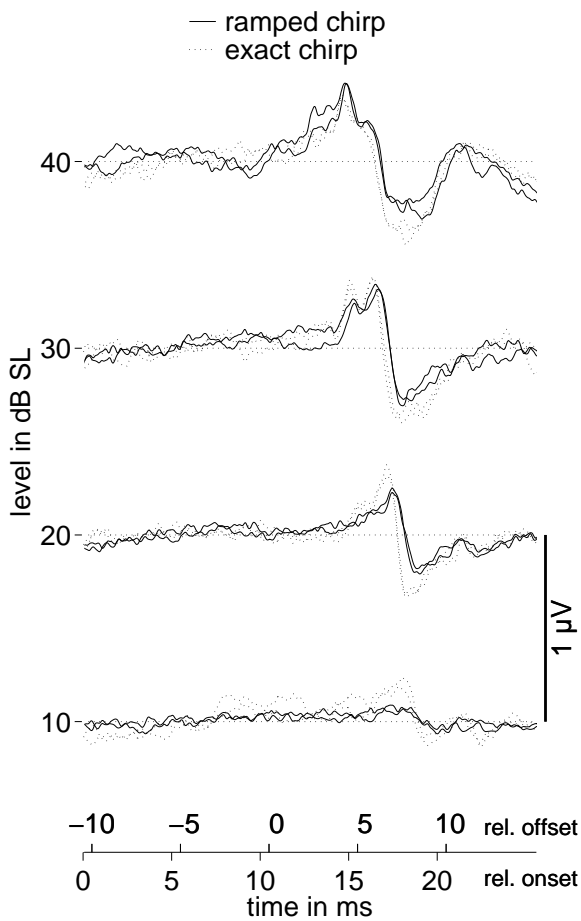
It could be argued that the relatively abrupt *offset* of the chirp is responsible for the generation of wave-V amplitude. Although such an argument would *not* explain the observation of an increased amplitude compared to the click response, a chirp stimulus was generated with sufficiently long ramps to preclude the possibility that purely onset- and offset effects are responsible for wave-V amplitude generation. A rise time of 3 ms and a fall time of 0.5 ms were applied.<sup>3</sup> Figure 2.5 shows the corresponding ABR recordings for subject CR (solid lines). In addition, the corresponding data with the exact chirp without ramps are replotted in the figure and indicated as dotted lines. In comparison with the original chirp without ramps, there is only a slight decrease in amplitude for this subject. This is most likely due to the attenuation of frequencies higher than about 6 kHz that normally also contribute to the generation of wave-V amplitude. Note, however, that the overall level of the ramped chirp had to be increased by 2 dB to yield the same sensation level.

Average data across subjects obtained with the ramped chirp are indicated in Fig. 2.3 (open diamonds). Like the original chirp without ramps, the ramped chirp elicits a significantly larger amplitude ( $p < 0.05$ ;  $N = 10$ ) than the click for the stimulation levels 20–40 dB SL. The difference is not significant for 10 dB SL.

All results taken together show that wave-V amplitude is increased when a rising broadband chirp is used instead of a click. This is the case although the duration of the chirp is about 10 ms, which is a factor of 125 longer than the click duration. This result is in contrast to the generally accepted view in the literature that the conventional ABR is an electrophysiological event only evoked by onset or offset of an acoustic stimulus.

---

<sup>3</sup>A larger fall time than 0.5 ms would also attenuate energy at medium frequencies. For example, if a fall time of 1 ms was used, frequencies around 2.5 kHz would be attenuated by about 3 dB. Higher frequencies would be attenuated up to 12 dB more than in case of the shorter ramp. We think that a 0.5-ms fall time and a 3-ms rise time for the chirp is more than a fair choice for the comparison of chirp and click efficiency in evoking ABR. In addition, in a preliminary study, we performed an experiment where the stimulus consisted of a *continuous* alternating sequence of chirps: each rising chirp was followed by the temporally reversed (falling) chirp. Results showed “peaked” response patterns whereby amplitude and (relative) latency of the peaks directly corresponded to those obtained with “single” rising chirps. The results further clearly demonstrate the importance of considering the effects of basilar-membrane traveling wave on the formation of the ABR, at least for wave V.

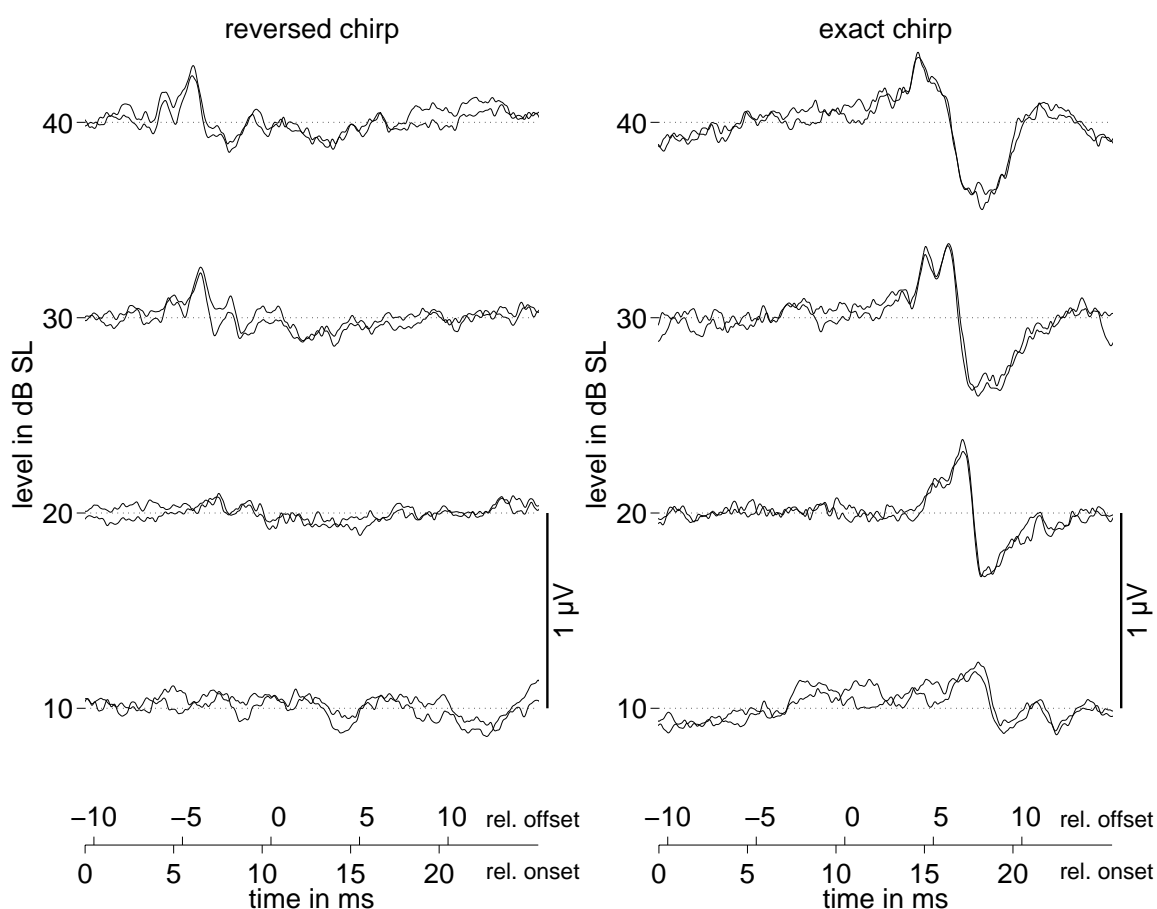


**Figure 2.5:** ABR from subject CR, evoked by the ramped broadband chirp. In addition, the potentials evoked by the original chirp without ramps are replotted from Fig. 2.2 and indicated as dotted curves.

#### 2.4.4 Effects of direction of frequency sweeping

If the argument holds that the “optimized” temporal course of the frequency sweeping is responsible for maximal synchronization, then a temporally *reversed* broadband chirp should yield a smaller response amplitude. The reversed chirp starts with high frequencies and sweeps nonlinearly in time towards low frequencies. The onset is therefore much steeper than that of the original chirp so that one should expect a larger response if ABR is determined by the steepness of the stimulus onset. The magnitude spectra of the reversed chirp and the original chirp, of course, are identical.

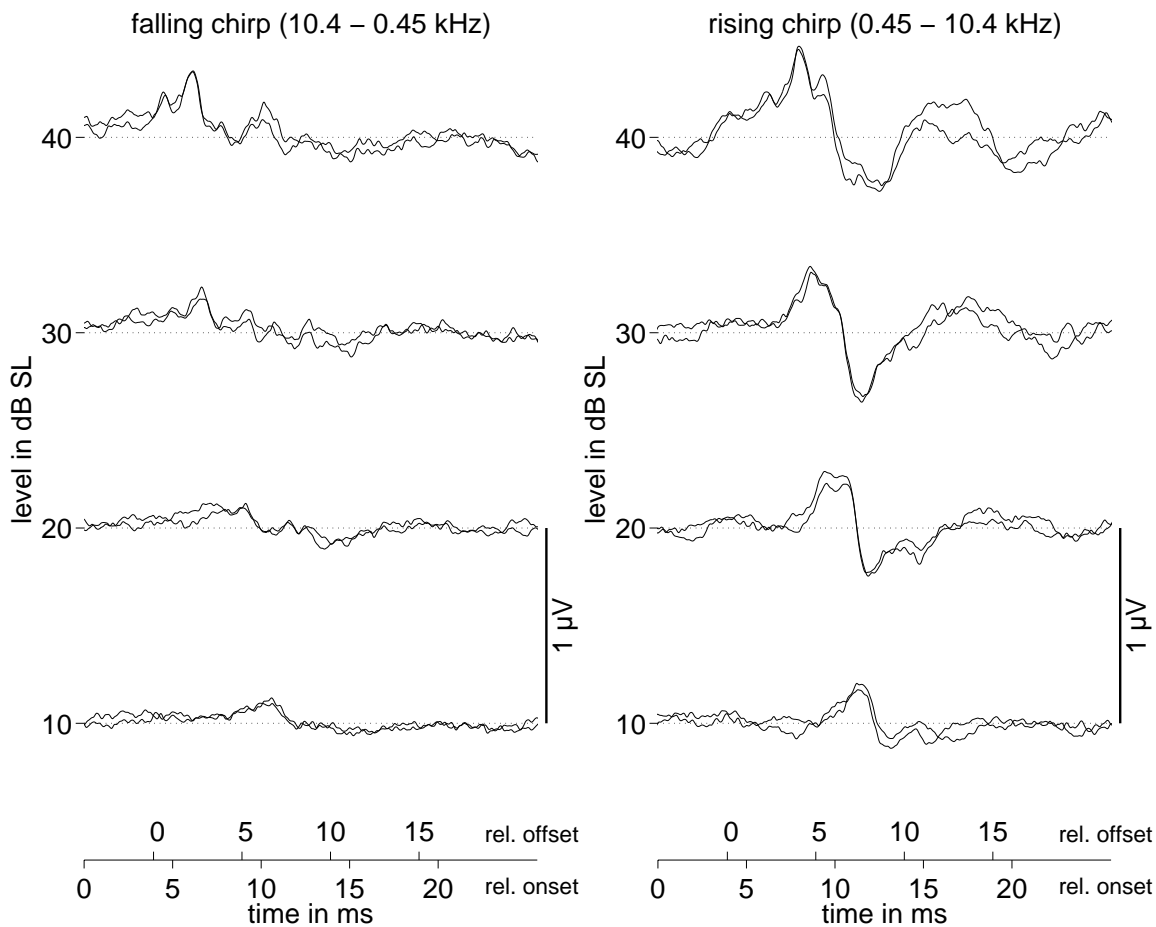
Figure 2.6 (left panel) shows the ABR for the reversed chirp (subject CR). For comparison, the right panel of Fig. 2.6 shows the ABR elicited by the rising chirp, for the same range of stimulation levels (10–40 dB SL), replotted from Fig. 2.2. It is apparent from the figure that in the case of the falling chirp, wave-V amplitudes are generally much smaller than those obtained with the rising chirp. The responses are also considerably smaller than those elicited by the click (see Fig. 2.2).



**Figure 2.6:** Left: ABR from subject CR, elicited by the temporally reversed broadband chirp (10.4–0.1 kHz). Parameters as in the previous figures. Right: ABR from the same subject, elicited by the rising broadband chirp (replot from Fig. 2.2, but only for the levels 10–40 dB SL).

The average data for the reversed chirp are indicated as filled downward-triangles in Fig. 2.3. Wave-V amplitude is significantly smaller ( $p < 0.05$ ;  $N = 10$ ) for the reversed chirp than for the rising chirp (filled boxes) for all stimulation levels 10–40 dB SL. The reversed-chirp amplitude was also significantly smaller than the click response (filled circles) for the levels 20–40 dB SL ( $p < 0.05$ ;  $N = 10$ ), while the difference was not significant for 10 dB SL.

However, because of the long duration of the chirp (10.52 ms), the response to the early (high-frequency) part may interfere with responses to the later (low-frequency) part of the stimulus. For this reason, a further experiment was run with a shorter chirp (3.92 ms), whose spectrum stretches from about 0.45 to 10.4 kHz (in contrast to 0.1 to 10.4 kHz as before). Figure 2.7 shows, for subject CR, the brainstem potentials evoked by the falling



**Figure 2.7:** ABR from subject CR, elicited by the 3.92 ms-chirp with a spectrum in the frequency range between 0.45 and 10.4 kHz. Left: temporally reversed (falling) chirp (10.4–0.45 kHz); right: rising chirp (0.45–10.4 kHz).

chirp (left panel) in comparison with the corresponding rising chirp (right panel). As in Fig. 2.6, the amplitude of wave V is much smaller in the case of stimulation with the falling chirp than with the rising chirp. Note that the responses evoked by the two reversed chirps (from Figs. 2.6 and 2.7) are almost identical, whereas the responses evoked by the two rising chirps differ to some extent. The average data for the 0.45–10.4 kHz chirp are indicated as open downward-triangles in Fig. 2.3. There is no significant difference in wave-V amplitude between the two chirps for 10 dB SL. However, for the levels 20–40 dB SL, the amplitude is significantly larger ( $p < 0.05$ ;  $N = 6$ ) for the 0.1–10.4 kHz chirp than for the 0.45–10.4 kHz chirp. This difference in wave-V amplitude directly reflects the contribution of the low-frequency components (100–450 Hz) to the ABR in the case of the rising chirp.

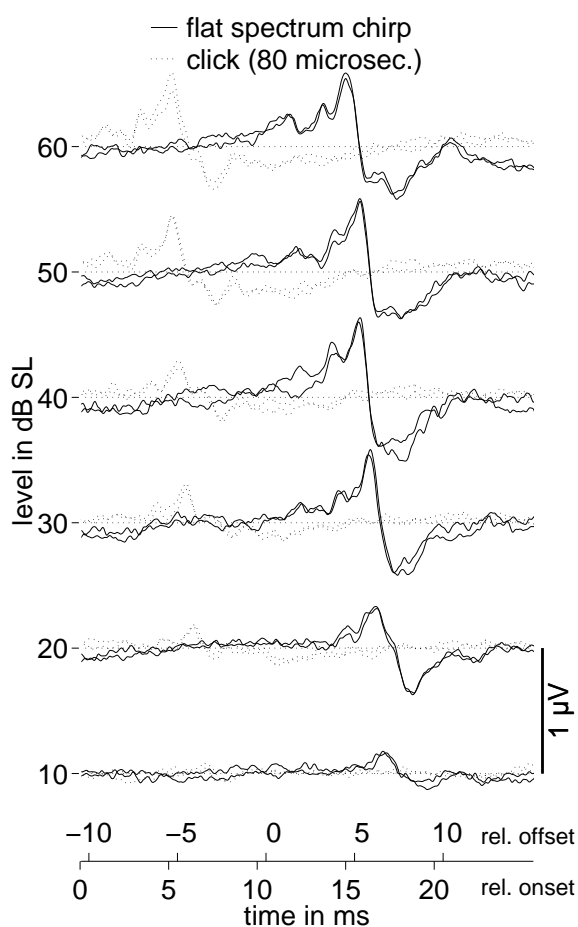
These results are compatible with the hypothesis that compensation of travel time differences across frequency causes an optimal synchronization, whereas the reversed stimulation leads to a less effective activation (although the onset of the reversed chirp is much steeper than that of the rising chirp). Falling sweeps probably produce sequential activation of high-frequency fibers followed by low-frequency fibers. This may lead to a desynchronized neural activation at the brainstem level, as implied by the results of Shore and Nuttall (1985) at the level of VIIIth nerve and cochlear nucleus (CN). In “far-field” recordings as considered in the present study, the observed effects may also be reflected by phase cancellation of the potentials due to superimposition of wave V from one frequency and wave V from another frequency (e.g., Scherg and von Cramon, 1985, see discussion).

### 2.4.5 Effects of spectral composition

It is not clear which spectral shape is optimal for broadband stimulation. It is also not clear, how and at which level, integration *across* frequency is realized in the auditory system. Hence, it may be argued that the observed differences between responses evoked by a click and a chirp stimulus is produced by their different spectral shape. To rule out this argument, ABR elicited by the flat-spectrum chirp from Fig. 2.1 (middle panel), with a flat amplitude spectrum corresponding to that of the click, are compared to click-evoked responses.

Figure 2.8 shows the corresponding ABR for this chirp for subject CR (solid curves). For direct comparison, the corresponding click-evoked responses for the same subject (replotted from Fig. 2.2, left panel) are indicated as dotted curves. Wave-V amplitudes are much larger for the chirp than for the click. The amplitudes are even larger than for the “normal” rising chirp without specific spectral weighting (Fig. 2.2, right panel), particularly at the highest stimulation levels. The average data for wave-V amplitude, obtained with the flat-spectrum chirp, are plotted as filled diamonds in Fig. 2.3. Wave-V amplitude is significantly larger ( $p < 0.05$ ;  $N = 10$ ) for this chirp than for click stimulation for *all* levels (10–60 dB SL). However, the difference in amplitude obtained with the flat-spectrum chirp and the original rising chirp (square symbols) was not statistically significant.

A similar approach has been earlier described in a study by Lütkenhöner *et al.* (1990). They also generated a rising chirp stimulus with a flat amplitude spectrum where - in contrast to the present study - the course of the instantaneous frequency was estimated from the relationship between the stimulus frequency and the experimentally obtained latency of the



**Figure 2.8:** ABR from subject CR elicited by the rising chirp with flat amplitude spectrum (from Fig. 2.1, middle panel). For direct comparison, click-evoked responses (dotted curves) for the same subject are replotted from Fig. 2.2.

corresponding frequency-specific ABR. The authors also found a larger wave-V amplitude with chirp than with click stimulation. However, the differences were smaller than those presented here, particularly at higher stimulation levels.

In summary, the presented data demonstrate that both dispersed timing as well as spectral composition of the stimulus strongly influence the potential pattern. The dispersed timing appears to be the dominant factor.

## 2.5 Discussion

### 2.5.1 Stimulus presentation: SL versus peSPL

In the present study, chirps and clicks were presented at the same sensation level. However, the corresponding peak-equivalent sound pressure level (peSPL) at the same sensation level, averaged across subjects, was about 12 dB *lower* for the chirp than for the click. For example,

at threshold (0 dB SL), the mean (and standard deviation) of the peSPL was  $45.9 \pm 3.7$  dB for the click while it was  $33.5 \pm 3.6$  dB for the chirp. It is unclear which measure is appropriate for ABR. Temporal integration of signal energy involved in behavioral threshold measures probably occurs at more central stages of auditory processing, and is most likely not reflected in ABR. Nevertheless, in many studies the stimuli are presented at the same hearing level (HL) to investigate potential amplitude in relation to the (normalized) average hearing threshold (0 dB HL). The strategy in the present study is thus very similar to the HL-measure, but hearing thresholds are determined individually. Whatever the proper calibration for ABR may be, if the responses obtained in the present study were plotted at the same peSPL instead of the same dB SL, the differences in the potential amplitude between chirp and click stimulation would be even larger.

### 2.5.2 Role of wave V behavior – spectral integration

By applying the derived response technique, [Don and Eggermont \(1978\)](#) revealed narrow-band contributions to the ABR from specific portions of the BM. They found that nearly the *whole* cochlear partition can contribute to the brainstem response. In their recordings, the amplitude behavior of wave V, as a function of the central frequency (CF) assigned to each narrow band, was different from waves I and III, depending upon the frequency range. Don and Eggermont found that for CFs below 2 kHz, the amplitudes for waves I and III drop rapidly as CF is decreased, whereas there is an increase in the amplitude of wave V. Therefore, at low CFs the only clear contribution to the ABR is to wave V. This indicates that the representation of cochlear activity in the various peaks probably is quite different for wave I and III on one hand, and wave V on the other hand ([Don and Eggermont, 1978](#)). Above 2 kHz, the wave-V behavior is the same as for the earlier waves. Thus, wave-V amplitude shows a flatter “frequency response” than the earlier waves and has an amplitude distribution which is nearly constant over the entire CF range.

By using a synchronizing chirp instead of a click as the stimulus, activity from *all* cochlear locations can contribute to the amplitude of wave V, which therefore is generally larger than that evoked by a click. This was demonstrated in the present study. [Don and Eggermont \(1978\)](#) stated that the discrepancy in the behavior of wave V with respect to the earlier

waves suggests some sort of neural reorganization at the level where wave V is generated.<sup>4</sup> The sharp initial positive potential is most likely generated by the lateral lemniscus as it enters in the inferior colliculus (IC), while the slow negative potential following this is likely a dendritic potential of the inferior colliculus (Hashimoto, 1982; Møller and Jannetta, 1982, 1986; Moore, 1987a,b). The central nucleus of the IC (ICC) is a purely auditory processing center; it is the main center for spatial auditory integration receiving most of the ascending information from auditory brainstem nuclei, and it has a curved laminar arrangement of cells, axons, and dendrites (e.g., Gummer and Zenner, 1996) which leads to an effective response. The convergence of pathways activates a large number of neurons in the IC; the wave-V potential therefore is of rather large amplitude, that obviously results from integration of activity from the whole range of auditory frequencies, and hence includes responses elicited by low-frequency stimulus components.

Interestingly, clear peaks corresponding to the earlier waves I–III could *not* be observed in the case of the original rising chirp (without specific spectral weighting) for any of the stimulation levels tested in the present study. This observation was true for all subjects, even if not investigated quantitatively. At the highest levels (50 and 60 dB SL), the early low-frequency energy of the chirp probably stimulates basal regions of the BM due to upward spread of excitation producing a response at about 8–9 ms after stimulus onset (as earlier mentioned in Sec. 2.4.1). However, it is not clear how these high-level responses are related to waves I–III. In contrast, we observed that, for the highest stimulation level of 60 dB SL, most of the subjects clearly showed the typical early peaks in their responses to the click as well as to the flat-spectrum chirp. In particular, the potential patterns at this level were very similar for these two stimuli for each subject. Thus, it appears that the spectral composition of the stimulus mainly determines the response pattern at high levels.<sup>5</sup>

---

<sup>4</sup>In a recent study, Don *et al.* (1997) presented a new derived-response measure where the wave-V amplitude of a stacked ABR was constructed by temporally aligning wave V of each derived-band ABR, and then summing the time-shifted responses. They found that the stacked response amplitude can detect small acoustic intracanalicular tumors in patients missed by standard ABR measures. It might be very interesting to compare the stacked wave-V amplitude obtained with the click with the chirp-evoked response amplitude obtained with the “standard” derived responses measure. Such experiments are currently in progress to specify the usefulness of the chirp for retrieving frequency-specific information.

<sup>5</sup>The orientation of the electrode configuration certainly plays a role for the shape of the potential pattern observed in the far-field. It has been shown that there are differences in click-evoked responses in horizontal



### 2.5.3 Assumption of linearity of BM characteristic

The chirp stimuli used in the present study were derived on the basis of a linear cochlea model (de Boer, 1980). It was assumed that the movements are so small that the fluid as well as the basilar membrane operate linearly. Over the normal range of hearing, the assumption of linearity may be well justified for the fluid (de Boer, 1980). With regard to the BM, however, its dynamics are certainly more complicated (e.g., Rhode, 1971; Ruggero, 1992). At low levels, BM dynamics may indeed be considered as nearly linearly. However, at higher levels, nonlinear cochlear mechanics complicate the responses to a frequency-changing signal. Ruggero and Rich (1983), for example, demonstrated that VIIIth-nerve fibers' phase response changes at high intensities resulting in two peaks which are 90° out of phase, instead of one peak commonly seen in period histograms. Since our interest in the present study was mainly focused on effects at levels between absolute threshold and about 40 dB SL, the assumption of linearity in the model calculations may be well justified. Also, the predictions of the linear model were primarily used for the correction of the dispersive behavior of the BM, which can be assumed to vary less with level than, e.g., the amplitude tuning characteristic for a certain frequency. Of course, a more general description for an extended level range would need to take nonlinear effects into account. The equations used so far should therefore be considered as a first-order approximation.

## 2.6 Summary and conclusions

A chirp stimulus was developed which theoretically produces synchronous discharges of VIIIth-nerve fibers along the length of the human cochlear partition. The equations defining the chirp were calculated to be the inverse of the delay-line characteristic of the cochlear partition on the basis of the linear cochlea model by de Boer (1980). The stimulus was tested for eliciting ABR. The underlying idea was to determine if units tuned to low CFs and vertical dipole orientations (e.g., Galbraith, 1994). Wave V is best seen in the vertical channel, while earlier components are well defined in the horizontal channel. Using a spatio-temporal dipole model, Scherg and von Cramon (1985) demonstrated a predominantly horizontal orientation underlying waves I–III. These horizontal dipoles appeared to reflect transverse propagation along the auditory nerve to the ipsilateral CN, and then via second order neurons crossing the midline to the contralateral superior olivary complex (SOC). Thus, it might be of interest to also investigate chirp-evoked ABR-recordings in horizontal dipole orientations.

(below 2 kHz) could be recruited synchronously into the brainstem response. It was shown in the present study that, in most level conditions, the chirp evokes a significantly larger wave-V amplitude than the click when presented at the same sensation level. This is the case although the duration of the chirp is about 10 ms, which is a factor of 125 longer than the click-duration used here. Since at the same sensation level, the peak-equivalent sound pressure level (peSPL) is about 12 dB *smaller* for the chirp than for the click, the difference in wave-V amplitude of the ABR recordings would be even larger if the stimuli would be presented at the same peSPL, or at the same peak-to-peak equivalent sound pressure level (ppeSPL). Thus, the conventional ABR should not be considered as an electrophysiological event purely evoked by the onset or offset of an acoustic stimulus. Instead, an appropriate temporal organization, determined by BM traveling wave properties, may increase neural synchrony at the level where wave V is generated. The temporally reversed chirp stimulus led to a smaller wave-V amplitude compared to the rising chirp and to the click. This may be due to desynchronized neural activation at the level where wave V is generated, as a result of sequential activation of high- followed by low-frequency fibers. Alternatively, the reduced potential amplitude may also result from cancellation in the “far-field” by superposition of wave V from one frequency on wave V from another frequency.

It was observed that not only temporal organization of the stimulus, but also its spectral shape, influences the ABR-pattern. The phase characteristic of the chirp, combined with a flat spectral distribution (as in case of the click), led to a large wave-V amplitude, but also to a more pronounced pattern of the earlier waves (at high stimulation levels), which is comparable with that evoked by the click. In contrast, responses evoked by the rising chirp without specific spectral weighting did not show clear earlier peaks I–III. This may be due to cancellation of overlapping responses at high stimulation levels where the early low-frequency energy in the chirp stimulates basal regions of the BM due to upward spread of excitation. Alternatively, or in addition, this may also be due to biased frequency representations at the level of the neural generators for waves I–III, while the generator for wave V probably has a flatter frequency response.

The use of the rising frequency chirp enables the inclusion of activity from lower frequency regions, whereas with a click or a falling chirp synchrony is decreased in accordance with decreasing traveling velocity in the apical region. The rising frequency chirp may therefore

be of clinical use in assessing the integrity of the entire peripheral organ, and not just its basal end.



## Chapter 3

# Frequency specificity of chirp-evoked auditory brainstem responses<sup>1</sup>

---

<sup>1</sup> This Chapter was published as a paper with the same title, written together with Torsten Dau, see [Wegner and Dau \(2002\)](#).

## Abstract

This study examines the usefulness of the upward chirp stimulus developed by Dau *et al.* [“Auditory brainstem responses with optimized chirp signals compensating basilar membrane dispersion,” *J. Acoust. Soc. Am.* **107**(3), 1530–1540 (2000)] for retrieving frequency-specific information. The chirp was designed to produce simultaneous displacement maxima along the cochlear partition by compensating for frequency-dependent traveling-time differences. In the first experiment, auditory brainstem responses (ABR) elicited by the click and the broadband chirp were obtained in the presence of high-pass masking noise, with cut-off frequencies of 0.5, 1, 2, 4 and 8 kHz. Results revealed a larger wave-V amplitude for chirp than for click stimulation in all masking conditions. Wave-V amplitude for the chirp increased continuously with increasing high-pass cut-off frequency while it remains nearly constant for the click for cut-off frequencies larger 1 kHz. The same two stimuli were tested in the presence of a notched-noise masker with one-octave wide spectral notches corresponding to the cut-off frequencies used in the first experiment. The recordings were compared with off-line calculated derived responses from the high-pass masking conditions. No significant difference in response amplitude between click and chirp stimulation was found for the notched-noise responses as well as for the derived responses. In the second experiment, responses were obtained using narrow-band stimuli. A low-frequency chirp and a 250-Hz tone pulse with comparable duration and magnitude spectrum were used as stimuli. The narrow-band chirp elicited a larger response amplitude than the tone pulse at low and medium stimulation levels. Overall, the results of the present study further demonstrate the importance of considering peripheral processing for the formation of ABR. The chirp might be of particular interest for assessing low-frequency information.

## 3.1 Introduction

A number of direct and indirect approaches have been used for retrieving frequency-specific information from the ABR. These approaches include different stimulus paradigms as well as different signal processing techniques. Stimulation with filtered clicks or different tone pulses is normally used, and selective masking techniques are generally employed. A limiting factor for eliciting frequency-specific ABR in the frequency region below 2 kHz is related to

cochlea mechanics and to the time-frequency uncertainty principle applied to the acoustic stimulus.

A straightforward approach to obtain frequency-specific ABR has been the stimulation by brief tone pulses with a short rise and fall time (e.g., [Kodera \*et al.\*, 1977](#); [Suzuki \*et al.\*, 1977](#); [Klein and Teas, 1978](#); [Coats \*et al.\*, 1979](#); [Purdy \*et al.\*, 1989](#); [Conijn \*et al.\*, 1993](#); [Beattie and Torre, 1997](#); [Bunke \*et al.\*, 1998](#)). As a compromise between frequency specificity and sufficient synchronization capability of the stimulus, [Davis \(1976\)](#) suggested the use of tone pulses with rise and fall times equal to two cycles of the stimulus frequency, and a plateau time equal to one cycle. It was found that high-frequency pulses (2 kHz and higher) elicit ABR which are similar to click-evoked responses (e.g., [Terkildsen \*et al.\*, 1975](#); [Gorga \*et al.\*, 1985](#); [Laukli and Mair, 1986](#); [Kileny, 1981](#); [Conijn \*et al.\*, 1992b](#); [van der Drift \*et al.\*, 1987](#)). Secondly, it has been shown that low-frequency pulses (below 2 kHz) of higher intensity, however, elicit ABR which include strong contributions originating from the more basal regions of the cochlea (e.g., [Beattie and Kennedy, 1992](#); [Gorga and Thornton, 1989](#)). On the other hand, responses evoked by less intense low-frequency tone pulses are difficult to identify since the larger rise time required to obtain a sufficient narrow bandwidth of the acoustic stimulus is not effective in synchronizing neural discharges ([Kramer and Teas, 1979](#); [Laukli and Mair, 1986](#); [Hoke \*et al.\*, 1991](#)). Therefore, it was argued that ABR elicited by stimulation with brief tone pulses of frequencies below about 2 kHz are only poor predictors of low-frequency behavioral thresholds ([Davis and Hirsh, 1976](#); [Debruyne, 1982](#); [Laukli, 1983a,b](#); [Laukli and Mair, 1986](#); [Laukli \*et al.\*, 1988](#); [Scherg and Volk, 1983](#); [Sohmer and Kinarti, 1984](#); [Weber, 1987](#)).

As a consequence, masking techniques have been suggested as an appropriate paradigm to obtain frequency-specific responses. The masker serves either to eliminate unwanted non frequency-specific contributions to the ABR by selectively masking regions of the cochlea which are outside the region to be stimulated; e.g., by notched-noise masking or high-pass noise masking ([Terkildsen \*et al.\*, 1975](#); [Picton \*et al.\*, 1979](#); [Stapells and Picton, 1981](#); [Pratt and Bleich, 1982](#); [Jacobson, 1983](#); [Stapells \*et al.\*, 1990](#); [Beattie \*et al.\*, 1992](#); [Beattie and Kennedy, 1992](#); [Conijn \*et al.\*, 1992a,b](#); [Abdala and Folsom, 1995a,b](#); [Oates and Stapells, 1997a](#)). Alternatively, the neural activity in specified cochlea regions can be selectively suppressed by computing off-line the difference waveform between the masked and unmasked responses, e.g., derived response technique ([Don and Eggermont, 1978](#); [Eggermont, 1976](#);

Eggermont and Don, 1980; Kramer, 1992; Noursak and Stapells, 1992; Donaldson and Ruth, 1993; Don *et al.*, 1994, 1997; Oates and Stapells, 1997b) or pure-tone masking (Klein and Mills, 1981a,b; Klein, 1983; Folsom, 1984, 1985; Pantev *et al.*, 1985; Mackersie *et al.*, 1993; Wu and Stapells, 1994).

Using the high-pass noise masking derived ABR technique, Don *et al.* (1994) investigated the effect of the temporal variability in the neural conduction time and the effect of variability in the cochlear response times on wave-V amplitude of the compounded ABR to clicks. They adjusted for differences in neural conduction time (I–V delay) through compression or expansion of the derived response times and adjusted for differences in the cochlear response times through (individual) shifts of the derived ABR waveforms. Compensation for the I–V variability had little effect while compensation for cochlear response times greatly affected the amplitude of wave V of the compounded ABR. The study demonstrated the powerful influence of the temporal aspects of cochlear activation and response times on the component amplitude of the compounded ABR.

Recently, Dau *et al.* (2000) developed an upward chirp stimulus that theoretically produces simultaneous displacement maxima by canceling traveling-time differences along the cochlear partition. The equations determining the temporal course of the chirp were derived on the basis of a cochlea model (de Boer, 1980) and were calculated to be the inverse of the delay-line characteristic of the human cochlea partition. The fundamental relationship between stimulus frequency and place of maximum displacement was taken from Greenwood (1990). ABR evoked by the broadband chirp showed a larger wave-V amplitude than click-evoked responses. Dau *et al.* (2000) demonstrated that the ABR is not an electrophysiological event purely evoked by onset or offset of an acoustic stimulus but that an appropriate temporal organization, determined by basilar-membrane (BM) traveling-wave properties, may significantly increase synchrony of neural discharges. The use of the upward broadband chirp enables the extension of activity to lower frequency regions whereas click synchrony is decreased in accordance with decreasing traveling velocity in the apical region of the cochlea.

The present paper examines the usefulness of this chirp for estimating frequency-specific information. In the first experiment, ABR evoked by the broadband chirp in the presence of high-pass and notched-noise masking are compared with corresponding click-evoked responses for the same subjects. The second experiment investigates ABR obtained with



narrow-band stimuli: responses elicited by a low-frequency chirp are compared with tone-pulse evoked responses whereby the chirp and the tone pulse were designed to have similar duration and magnitude spectrum so that they mainly differ in their phase characteristic. The role of cochlear processing for brainstem responses and the possible application of the chirp for assessing low-frequency information are considered.

## 3.2 Method

### 3.2.1 Subjects

Nine normal hearing subjects (one female and eight male) with no history of hearing problems and audiometric thresholds of 15 dB HL or better participated in the experiments. All subjects were between 24 and 36 years of age, and either volunteered or were paid for the experiment.

### 3.2.2 Apparatus

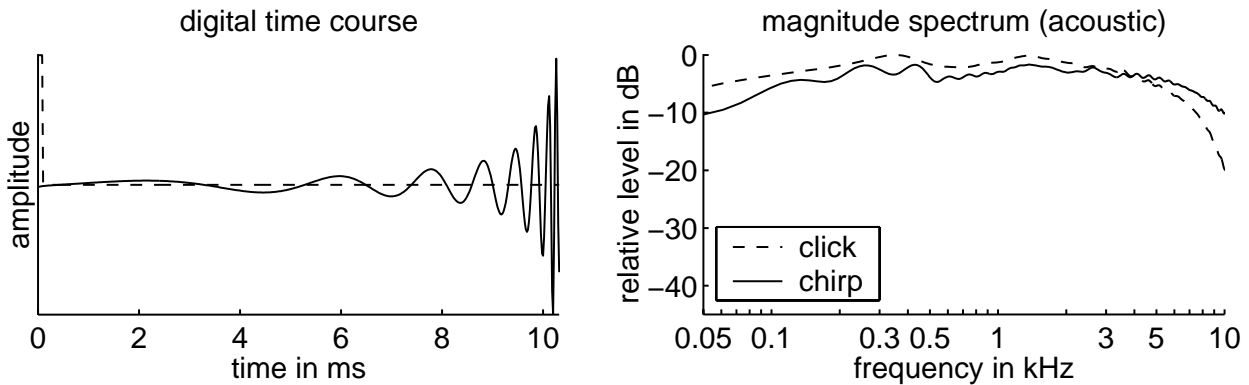
The experiments were carried out with a PC-based computer system which controlled stimulus presentation and recording of evoked potentials. A DSP-card (Ariel DSP32C) converted the digitally generated stimulus (25 kHz, 16 bit) to an analogous waveform.

The masking noise in the two experiments was generated by feeding broadband white noise from a random noise generator (TDT WG2) to two cascaded filters (TDT PF1). The output of the second filter was attenuated (TDT PA4) and then added to the stimulus by a signal mixer (TDT SM3). The output of the signal mixer was connected to a digitally controlled audiometric amplifier, which presented the stimulus through an insert earphone (Etymotic Research ER-2) to the subject.

Electroencephalic activity was recorded from the scalp via silver/silver chloride electrodes, attached to the vertex (positive) and the ipsilateral mastoid (negative). The forehead served as the site for the ground electrode. Inter-electrode impedance was maintained below 5 k $\Omega$ . Responses were amplified (80 dB) and filtered (95–1640 Hz, 6 dB/Oct.) with a commercially available ABR preamplifier (Hortmann Neurootometrie).<sup>2</sup> Extra amplification

---

<sup>2</sup>In the official data sheet of the preamplifier, a “hard-wired” high-pass cut-off frequency of 30 Hz is given. Unfortunately, we could not replicate this value and found a 3-dB cut-off of 95 Hz. The problem is that this



**Figure 3.1:** Temporal course (left panel) and acoustic spectra (right panel) of the chirp (solid line) and the click (dashed line) used in experiment 1. The chirp was defined in Dau et al. (2000) as “flat-spectrum chirp”. Its acoustic spectrum is similar to that of the click stimulus.

(Kemo VBF/40) was used to reach the optimum range for the A/D-converter. This amplification was in the range from 10 to 16 dB, resulting in a total amplification of 90–96 dB. The amplified signal was digitized by the DSP-card (25 kHz, 16 bit), which also performed artifact rejection and signal averaging. Responses were recorded for 40 ms following the stimulus onset.

### 3.2.3 Stimuli and procedure

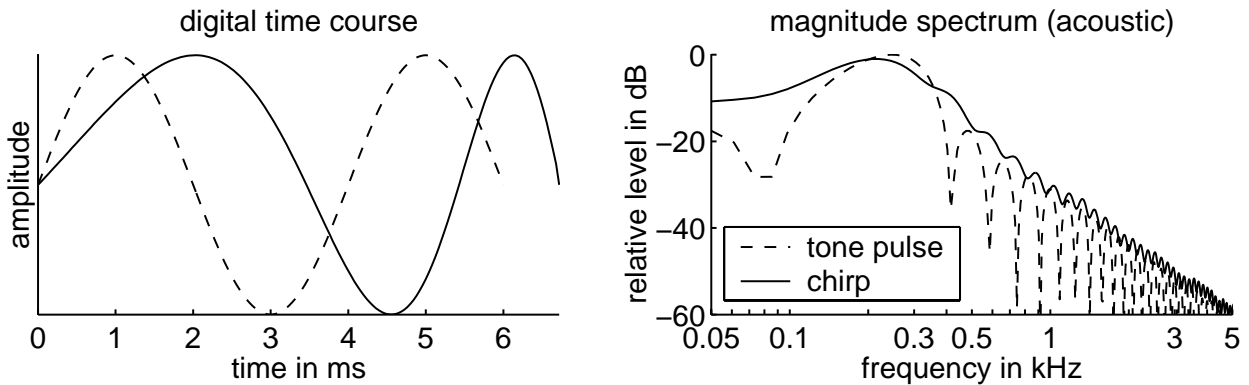
All chirp stimuli used in the present study were generated on the basis of the equations described in Dau et al. (2000). In the first two experiments, a broadband chirp was used with a magnitude spectrum corresponding to that of the click. This chirp thus represents the “flat-spectrum” chirp as defined in Dau et al. (2000). Its nominal edge-frequencies are 0.1 and 10.4 kHz, resulting in a duration of 10.48 ms. The chirp started and ended in zero phase and no windowing was applied. Figure 3.1 (left panel) shows the digital waveform of the chirp (solid curve). The corresponding acoustic spectrum is given in the right panel of the figure. Waveform and corresponding acoustic spectrum of the 80  $\mu$ s click are indicated as dashed curves in Fig. 3.1. The spectra were obtained by coupling the ER-2 insert earphone to a Brüel and Kjær ear simulator (type 4157) with a 1/2-in. condenser microphone (type 4134), setting will cut out a substantial portion of the wave-V amplitude which results in smaller responses overall, particularly for responses from lower frequency stimulus energy.

a 2669 preamplifier, and a 2610 measuring amplifier. The spectra were derived from fast Fourier transforms (FFTs) of 100-trial time-domain averages of the stimulus over an analysis time of 64 ms using a sampling rate of 25 kHz (Stanford Research System SR780). The waveforms were not windowed prior to FFT. In the recordings, both stimuli were presented at a stimulation level of 40 dB HL, corresponding to a peak-equivalent sound pressure level of 87 dB for the click and 80 dB for the chirp stimulus. This 7-dB difference between click and chirp sound pressure level reflects temporal integration of signal energy involved in behavioral threshold measures that probably occurs at more central stages of auditory processing and is most likely not reflected in ABR. The level of only 40 dB HL was chosen for two reasons: (i) chirp and click still exhibit a clearly detectable wave V (Dau *et al.*, 2000) when presented without noise masking, and (ii) the overall level and loudness of the stimuli in combination with additional high-pass or notched-noise maskers remain within a comfortable range.

In the second part of the study, ABR were obtained using a low-frequency chirp with nominal edge frequencies of 100 and 480 Hz. Results were compared with ABR obtained with a 250-Hz tone pulse. Figure 3.2 shows the digital temporal course (left panel) and the corresponding acoustic magnitude spectrum (right panel) of the low-frequency chirp (solid curve) and the tone pulse (dashed curve). Both stimuli were designed to exhibit three “half waves” beginning and ending with zero. The tone-pulse frequency was chosen in such a way that duration and magnitude spectrum were similar to those of the low-frequency chirp. As can be seen in the right panel of Fig. 3.2, the magnitude spectrum of the tone pulse is slightly narrower than that of the low-frequency chirp with the largest differences occurring at frequencies below about 100 Hz. The stimulation level was varied between 20 and 40 dB HL, in 5-dB steps. The peak-equivalent sound pressure level at hearing threshold (0 dB HL) was 40 dB for the tone pulse and 42 dB for the low-frequency chirp.

To determine the hearing level for the different stimuli, the absolute hearing thresholds were measured individually with an adaptive alternative forced choice (3-AFC) procedure. The average over all subjects in the present study were considered as representing 0 dB HL.

The subject lay on a couch in an electrically shielded, soundproof room, and electrodes were attached. The subject was instructed to keep movement at a minimum, and to sleep if possible. The lights were turned out at the beginning of the session. Each session lasted between one and two hours, depending on the subject’s ability to remain still. The ear of stimulation was chosen randomly, i.e., for each subject one ear was chosen and then main-



**Figure 3.2:** Temporal course (left panel) and acoustic spectra (right panel) of the low-frequency stimuli used in experiment 2. The narrow-band chirp (solid curve) and the 250-Hz tone pulse (dashed curve) were designed to exhibit three “half waves” beginning and ending with zeros. The tone-pulse frequency was chosen in such a way that duration and magnitude spectrum were similar to those of the chirp.

tained. The acoustic signals were delivered at a repetition rate of 20 Hz for all stimulus conditions. A temporal jitter of  $\pm 2$  ms was introduced to minimize response superimposition from preceding stimuli. Thus the resulting inter-stimulus interval (ISI) was equally distributed between 48 and 52 ms. Each trial consisted of 3000 averages. For each stimulus condition, two independent trials were stored in separate buffers. These are illustrated as superimposed waveforms in the figures to show response replicability.

### 3.2.4 Experimental masking paradigms

Two different experimental masking paradigms were used in the first part of the study for retrieving frequency-specific activity: derived responses obtained with the procedure proposed by [Don and Eggermont \(1978\)](#) and responses obtained with the notched-noise masking method. Responses to the broadband chirp were compared with corresponding click-evoked responses. In the present study, white noise was used as the masker which is different from the study of Don and Eggermont where pink noise was used.<sup>3</sup> In a first step, the level

<sup>3</sup>It is not clear which noise type is more appropriate for the analysis of frequency-specific contributions. Pink noise with its 3-dB reduction of energy per octave excites the cochlea roughly uniformly on a logarithmic frequency scale while white noise may lead to some overmasking at high frequencies and undermasking at low frequencies. On the other hand, the two stimuli used in the present study, the click and the chirp, have

of unfiltered broadband noise which was sufficient to obliterate the brainstem response was determined and this “masked” activity was recorded (Purdy *et al.*, 1989; Conijn *et al.*, 1990, 1992a). This was done for a stimulus level of 40 dB HL. Without changing any attenuation levels, the brainstem responses were recorded with this noise filtered in the following way: in case of the derived-response method, the noise was high-pass filtered at values in the order 0.5, 1, 2, 4, and 8 kHz. Then, including the unmasked (no noise) and the completely masked responses, a total of 7 recordings were obtained and stored. By successively subtracting response waveforms obtained in noise with high-pass cut-off frequencies separated by one octave, the narrow-band contributions to the ABR were obtained off-line, as suggested by Don and Eggermont (1978). In case of the notched-noise method, the spectral notches of the noise were represented by the octave-wide regions at 0.5–1 kHz, 1–2 kHz, 2–4 kHz, and 4–8 kHz, respectively. These notches represent those spectral regions from which stimulus-evoked activity can effectively contribute to the recorded ABR. The spectrum level of both high-pass noise and notched noise was nearly the same in the two experimental conditions (click: 32.6 dB, chirp: 32.5 dB).

In the second part of the study, responses evoked by the low-frequency chirp and the tone pulse were obtained. In order to ensure that neurons from more basal portions of the cochlea do not contribute to the evoked response, an additional set of recordings was obtained for the same stimuli with additional high-pass noise masking. The level of the unfiltered broadband noise which was sufficient to obliterate the brainstem response of the 40-dB HL signal was determined in a first step. Recordings were obtained using the noise high-pass filtered at 1 kHz without changing any attenuation levels. For the remaining signal levels it was assumed that the signal-to-noise ratio at “masked threshold” remains the same. Such a strategy has also been used by other investigators (e.g., Conijn *et al.*, 1990, 1992a).

### 3.2.5 Statistical analysis

Wave-V peak-to-peak amplitude was analyzed in all stimulus conditions. The amplitude was measured from the peak to the largest negativity following it. For each condition, wave-V amplitude was averaged across subjects. A Wilcoxon matched-pairs signed-rank test

---

a flat frequency spectrum and thus themselves do not excite the cochlea uniformly suggesting white noise as an appropriate masker. Also, white noise has been used in most of the notched-noise masking studies in the literature.

( $\alpha = 0.05$ ) was performed to verify whether the response amplitude differed significantly for the two comparison stimuli. Throughout the present paper, responses are shown for exemplary subjects. Mean data for wave-V amplitude, averaged across the nine subjects, are summarized in additional figures.

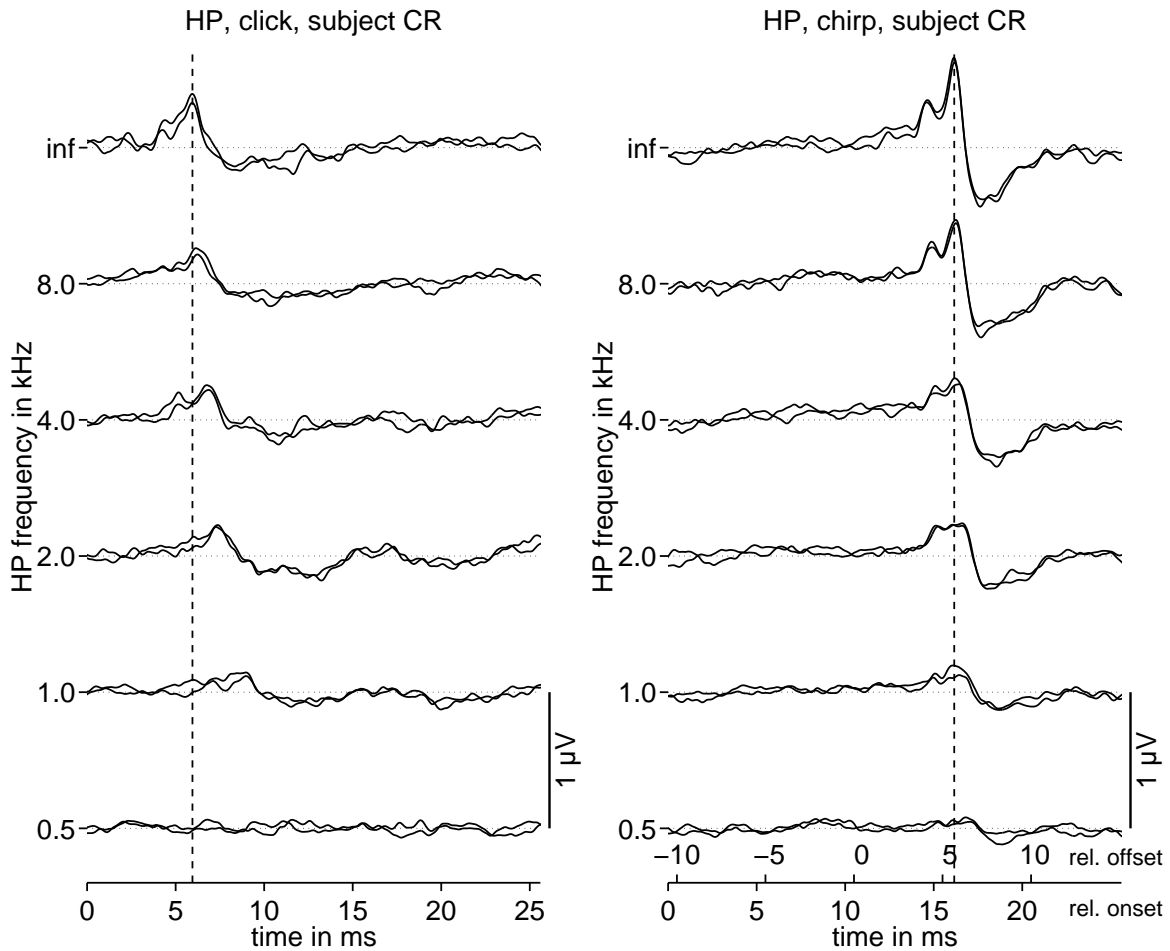
## 3.3 Results

### 3.3.1 Click- versus chirp-evoked responses using noise masking

The left panel of Fig. 3.3 shows, for subject CR, a series of brainstem responses to the click obtained in the presence of high-pass noise. As a reference, the unmasked response is shown as top curve (indicated as “inf”) in the figure. Wave V is the only clear peak in the unmasked as well as in most of the noise masking conditions. However, wave-V amplitude is strongly reduced or absent for noise cut-off frequencies below 2 kHz. These observations are consistent with the results of [Don and Eggermont \(1978\)](#) apart from the fact that their potential waveforms exhibited a larger wave-V amplitude and more distinct earlier waves I and III for the unmasked response caused by the higher stimulation level they used (60 dB SL instead of 40 dB in the present study). The dashed vertical line in Fig. 3.3 indicates wave-V latency for the unmasked response. As expected, there is a gradual increase in latency for the wave-V peaks as the high-pass cut-off frequency decreases, since the non-masked contributions to the ABR are located in more and more apical portions of the cochlea (e.g., [Laukli, 1983b](#); [Burkard and Hecox, 1983](#); [Don et al., 1994](#)).

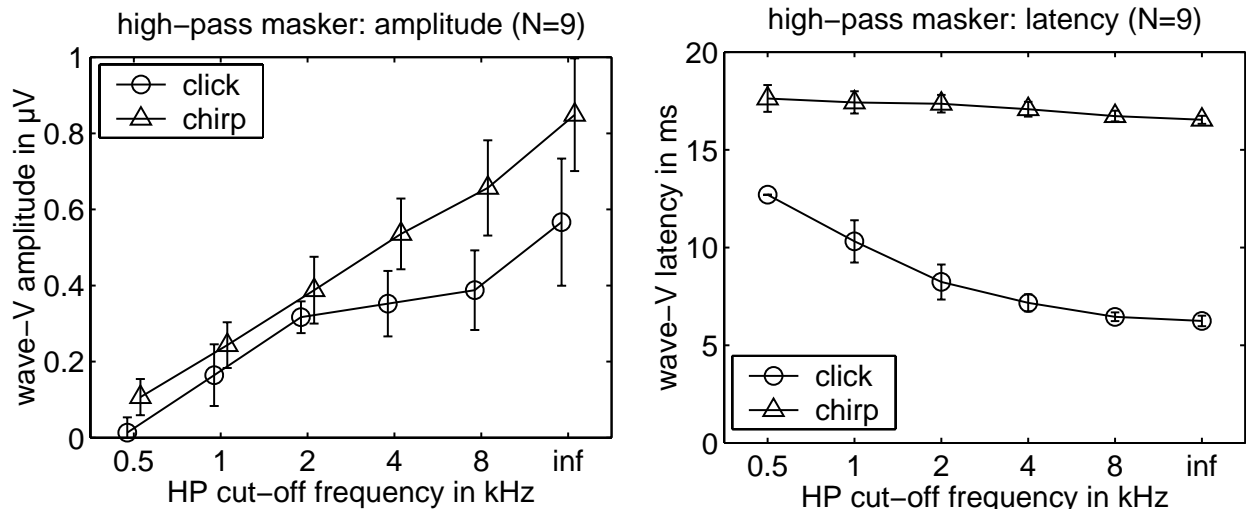
The right panel of Fig. 3.3 shows the corresponding potential patterns evoked by the broadband chirp. Wave-V amplitude is clearly detectable in all conditions and larger than the click-evoked amplitude. The earlier waves (particularly wave III) are visible in the unmasked condition (top curve) and for the 8-kHz masking condition. Wave-V latency is roughly constant across all conditions. This latter finding was expected because of the specific phase characteristic of the chirp stimulus which was designed to compensate for travel-time differences on the BM ([Dau et al., 2000](#)).

Figure 3.4 shows the mean results for click and chirp stimulation, averaged across the nine subjects. The left panel shows wave-V amplitude for the click (circles) and for the chirp (triangles), as a function of the cut-off frequency of the noise masker. For the click, wave-V



**Figure 3.3:** Individual ABR evoked by the click (left panel) and the chirp (right panel) in the presence of high-pass noise masking. Parameter is the cut-off frequency of the noise. The unmasked responses (“inf”) are plotted on the top. To show response variability, the waveforms of the two independent buffers (3000 averages each) are shown on top of each other. The stimulus level was 40 dB HL. The stimulus presentation rate was 20/s. Subject: CR.

amplitude increases with increasing cut-off frequency up to 2 kHz but tends to saturate for higher frequencies. In contrast, for the chirp wave-V amplitude increases continuously with increasing cut-off frequency. The chirp-evoked wave-V amplitude is significantly larger than the click-evoked one in all conditions ( $N = 9$ ,  $\alpha = 0.05$ ). The right panel of Fig. 3.4 shows the mean values for wave-V latency. It can be seen that for the click (circles) the latency decreases with increasing masker cut-off frequency while it remains roughly constant for the chirp (triangles).

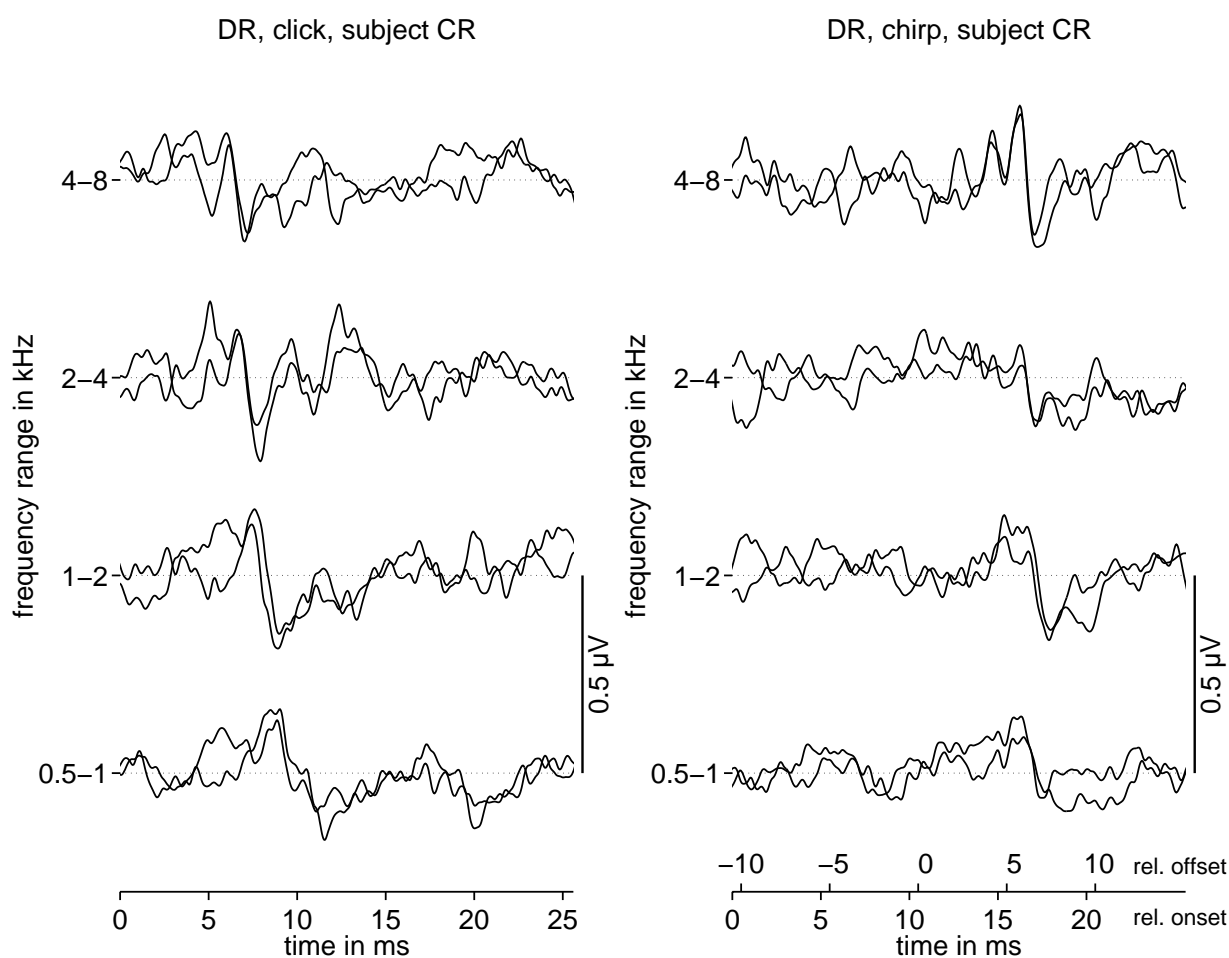


**Figure 3.4:** Mean wave-V amplitude (left panel) and latency (right panel) for click- and chirp stimulation, as a function of the high-pass noise cut-off frequency (from Fig. 3.3). Error bars indicate the standard deviation.

Figure 3.5 shows the derived responses for subject CR, obtained by successively subtracting the responses from Fig. 3.3. The left panel shows results for the click and the middle panel represents corresponding results for the chirp. Even though wave-V amplitude of all responses from the high-pass masking conditions (from Fig. 3.3) were larger for the chirp than for the click, the amplitude of the largest component of the derived responses in Fig. 3.5 is similar for the two stimuli, or even larger for the click. The mean data for wave-V amplitude are shown in Fig. 3.7 (left panel). Statistical analysis revealed no significant difference between the response amplitudes obtained with click and chirp.

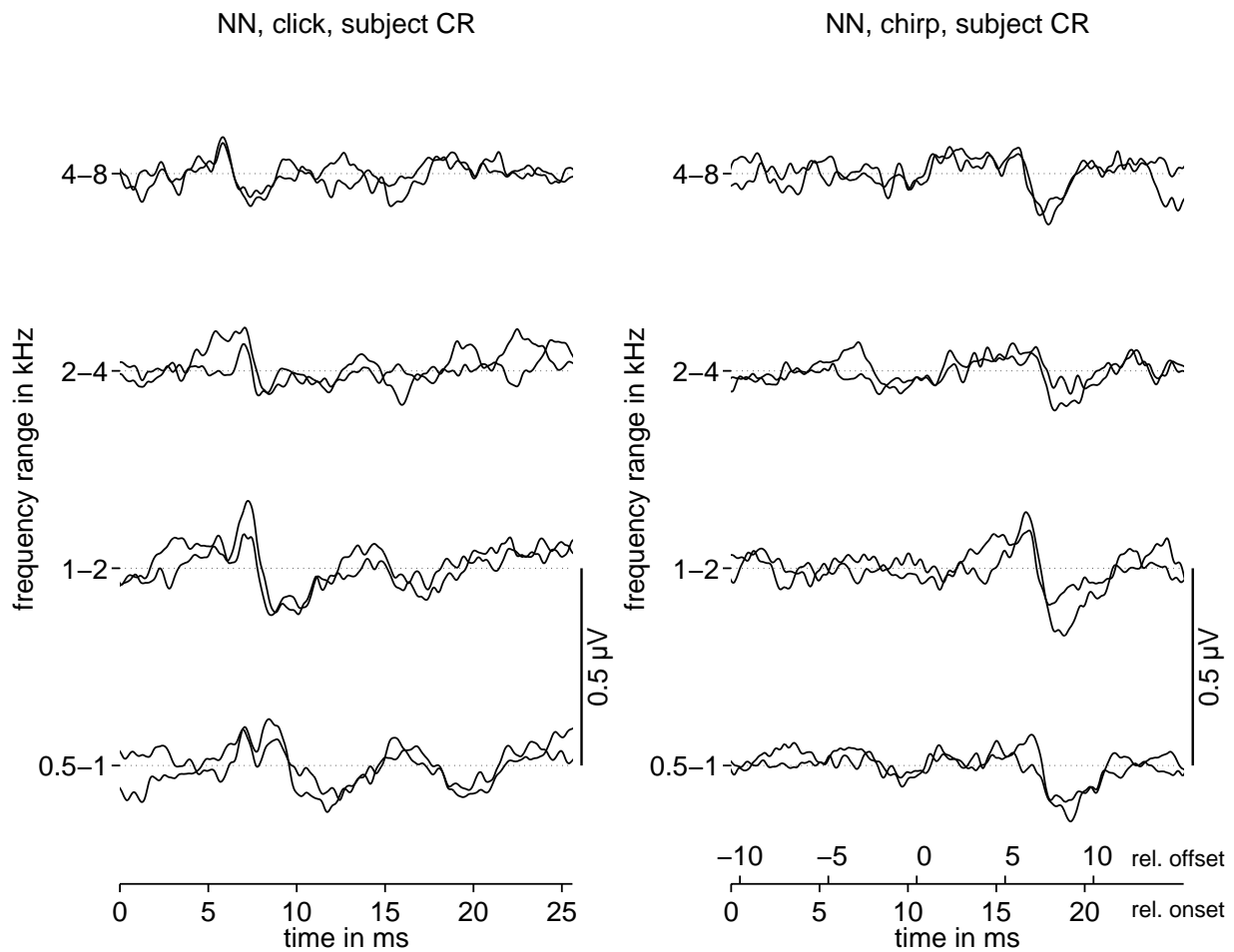
Finally, Fig. 3.6 shows click- and chirp-evoked responses obtained in the presence of a notched-noise (NN) masker. The response amplitudes are smaller than for the derived responses. This is most likely a consequence of the spread of excitation into the notch especially from the low-frequency noise band of the masker. The low-frequency part of the noise reduces the effective depth of the notch which is a much larger effect than the downward spread from the upper band (Picton *et al.*, 1979; Abdala and Folsom, 1995a; Beattie *et al.*, 1996). Wave-V amplitude is similar for the two stimuli in all NN conditions. The mean wave-V amplitude is plotted in the right panel of Fig. 3.7, as a function of the notch region of the noise masker. Wave-V amplitudes obtained with click and chirp stimulation do not differ significantly from one another.





**Figure 3.5:** Individual derived responses for click (left panel) and chirp stimulation (right panel). The responses were obtained by subtracting the high-pass masked ABR from the one shown immediately above it in Fig. 3.3.

In summary, larger response amplitudes were obtained for chirp than for click stimulation in the unmasked condition as well as in *all* high-pass masking conditions. However, no significant difference in response amplitude between click- and chirp- stimulation could be observed for the derived responses as well as for the notched-noise responses. This indicates, that even though neural synchrony is higher for the chirp than for the click, the contributions from one-octave wide frequency regions are not sufficient to obtain a significant difference in the far field.

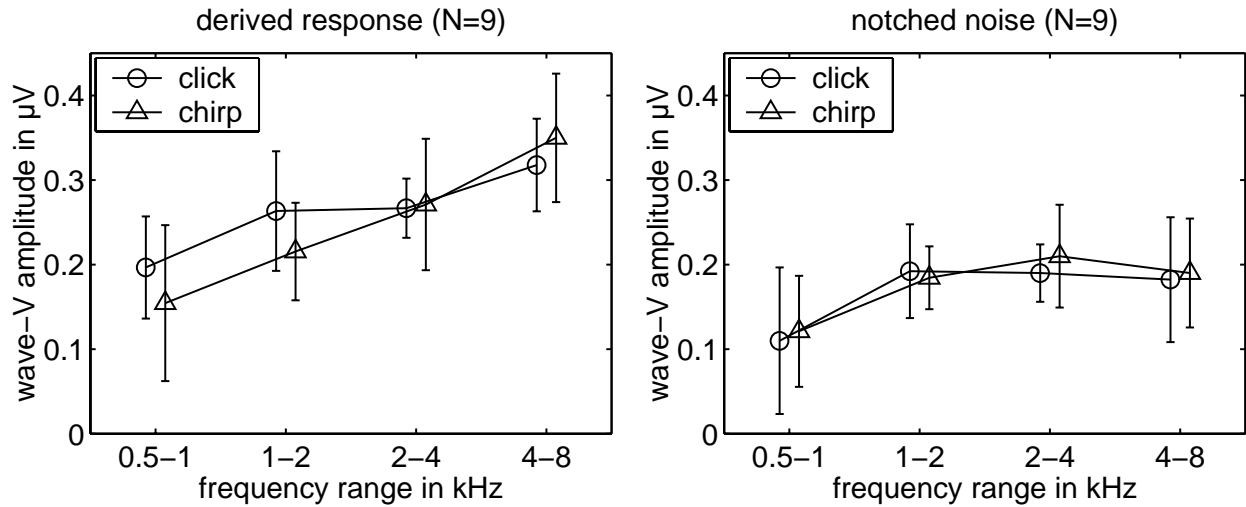


**Figure 3.6:** Individual ABR to clicks (left panel) and chirps (right panel) in the presence of a notched-noise masker. The spectral notches of the noise were one octave wide in each condition.

### 3.3.2 Tone-pulse versus low-frequency chirp-evoked responses

A more direct approach to retrieve frequency-specific ABR than the recording of click- or (broadband) chirp-evoked responses in combination with masking noise may be the use of brief tonal stimuli. In the present study, tone-pulse evoked responses were compared with responses elicited by a low-frequency chirp (see Fig. 3.2).

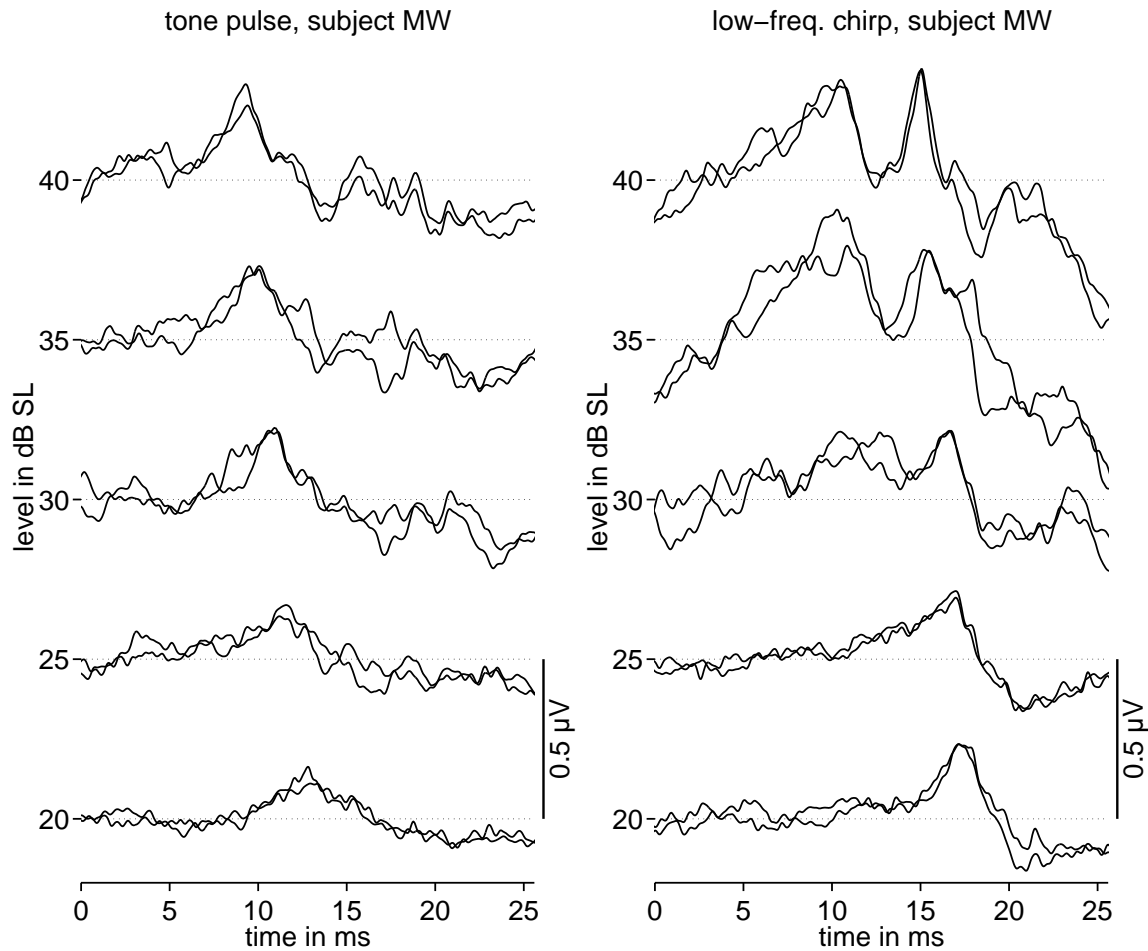
Figure 3.8 shows individual results for the 250-Hz tone pulse (left panel) and the chirp (right panel). First, for all stimulation levels tested, the chirp elicits a larger wave-V amplitude than does the tone pulse. Second, wave-V latencies of the chirp-evoked responses are shifted by about 5 ms towards larger values relative to those obtained with the tone pulse.



**Figure 3.7:** Mean amplitudes for the click- and chirp-evoked derived responses (left panel), and for notched-noise masking responses (right panel) (from Figs. 3.5 and 3.6). Error bars indicate standard deviation.

The responses for this particular subject have a larger amplitude than those for the other subjects but reflect the main trends observed in the mean data shown in Fig. 3.9. The left panel indicates wave-V amplitude as a function of the stimulation level. The chirp leads to a significantly larger wave-V amplitude than does the tone pulse, for the three lowest levels ( $N = 9$ ,  $\alpha = 0.05$ ). The right panel shows corresponding mean values for wave-V latency. For both stimuli, wave-V latency decreases with increasing level by about 2.6 ms consistent with findings from other studies (e.g., Gorga *et al.*, 1988). More importantly, there is a nearly constant latency shift of 5.1 ms between tone-pulse and chirp stimulation. The reason for the latency difference is the same as for the observed difference in the responses to click and broadband chirp if plotted relative to the stimulus onset (Dau *et al.*, 2000): by successively stimulating lower and higher frequency components with a sweeping rate determined by cochlear travel-time properties, activity should be more dominated near the stimulus end for the chirp. In contrast, for the tone pulse, activity should be more dominated near the beginning of the stimulus. This will be discussed in more detail in Sec. 3.4.

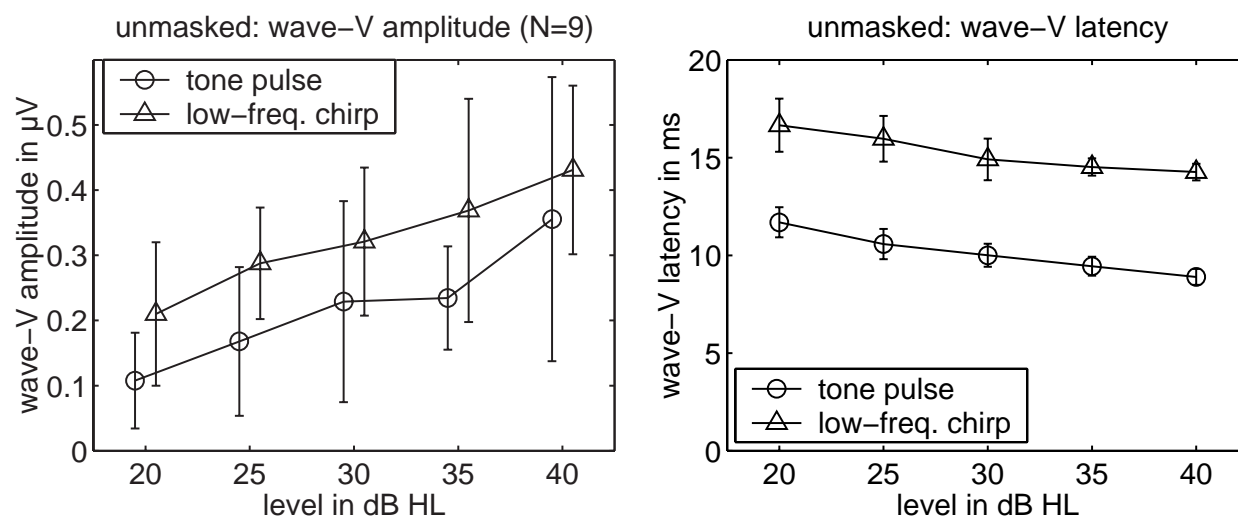
In fact, because of the short duration of the stimuli and because no ramps were used, also neurons tuned to medium and high frequencies may have contributed to the response, particularly at the higher stimulation levels. For this reason, additional responses with a high-pass noise masker (at a cut-off frequency of 1 kHz) were recorded. Individual results



**Figure 3.8:** Individual ABR to the 250-Hz tone pulse (left panel) and the low-frequency chirp (right panel). The stimulation level varied from 20 to 40 dB HL in 5-dB steps. Subject: MW

are shown in Fig. 3.10, and mean data are represented in Fig. 3.11. The response amplitude is markedly reduced compared to the no-noise condition from the previous experiment indicating that frequency components higher than about 1 kHz also contributed to the responses shown in Fig. 3.8. However, as for the unmasked conditions, statistical analysis revealed that for the three lowest stimulation levels the chirp causes a significantly higher wave-V amplitude than does the tone pulse.

These results demonstrate, that even a slight change in the phase characteristic of a stimulus can cause significant differences in the corresponding neural excitation. The low-frequency chirp leads to a higher synchronization of neural discharges which is also main-



**Figure 3.9:** Mean wave-V amplitude (left panel) and latency (right panel) for tone-pulse and low-frequency chirp stimulation. Error bars indicate standard deviation.

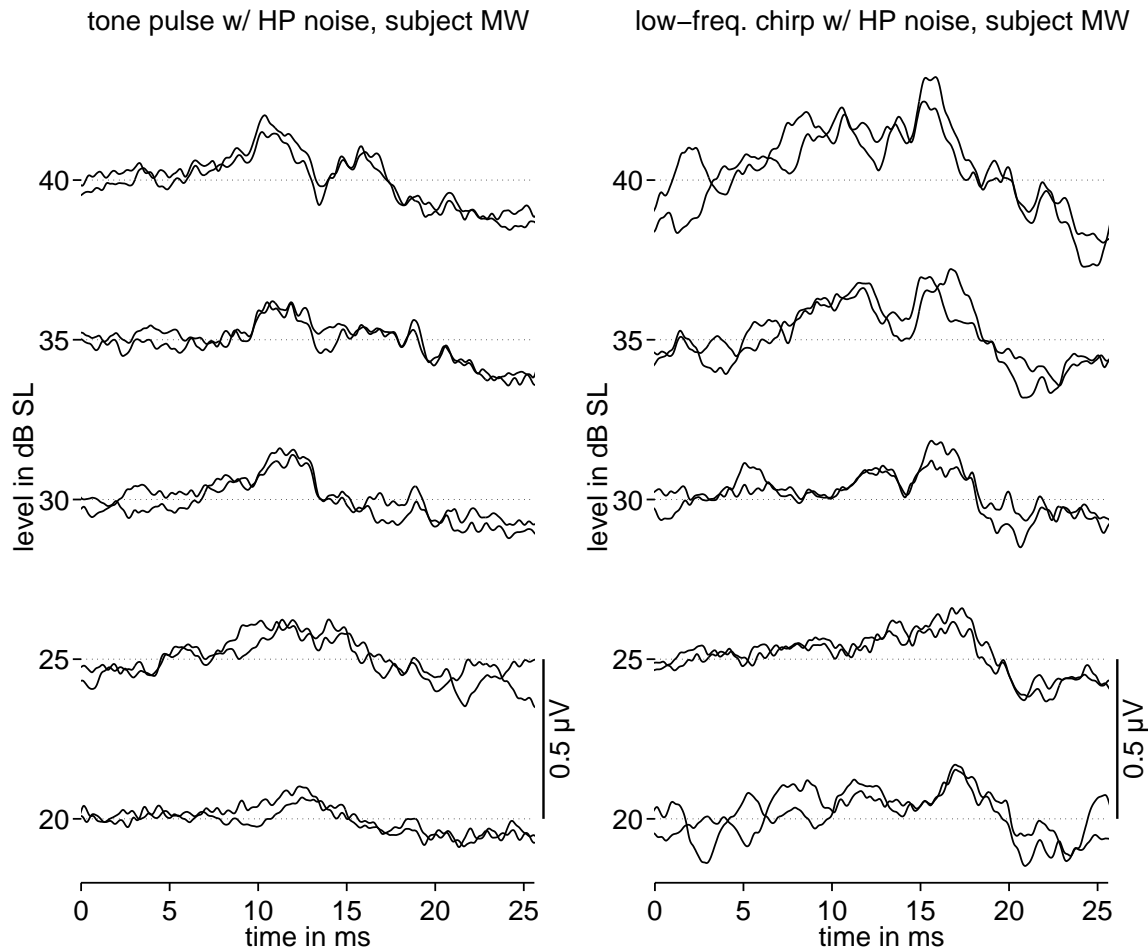
tained at higher than peripheral stages of signal processing, and is reflected in a larger value of wave-V amplitude.

However, since the evoked potential amplitude is affected not only by the degree of synchronization of neural activity but also by the number of neural elements involved, one may argue that the differences in response amplitude are due to differences in the overall neural excitation of the two stimuli. As was shown in Fig. 3.2, tone pulse and low-frequency chirp differ somewhat in their magnitude spectrum particular at frequencies below about 100 Hz which will lead to differences in their neural excitation pattern. This will be discussed in Sec. 3.4.2.

## 3.4 Discussion

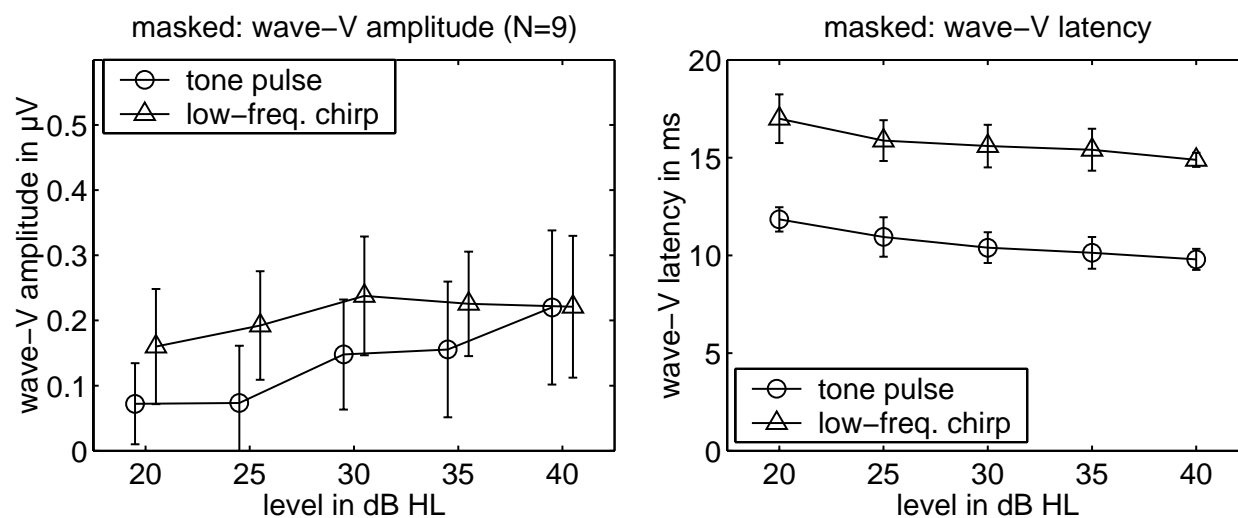
### 3.4.1 Broadband chirp versus click

In [Dau et al. \(2000\)](#) it was shown that, without any masking noise, the flat-spectrum chirp elicits a larger wave-V amplitude than the click for a large range of stimulation levels. In the present study, only one signal level (40 dB HL) was used. The results for the no-noise condition correspond to those in [Dau et al. \(2000\)](#). In order to investigate frequency-specific differences of ABR obtained with click and chirp stimulation, the stimuli of the present study



**Figure 3.10:** Individual ABR to the 250-Hz tone-pulse (left panel) and the low-frequency chirp (right panel). Stimulation as in Fig. 3.8 but with additional high-pass noise ( $f_c = 1$  kHz) which masks contributions from high-frequency portions of the basilar membrane.

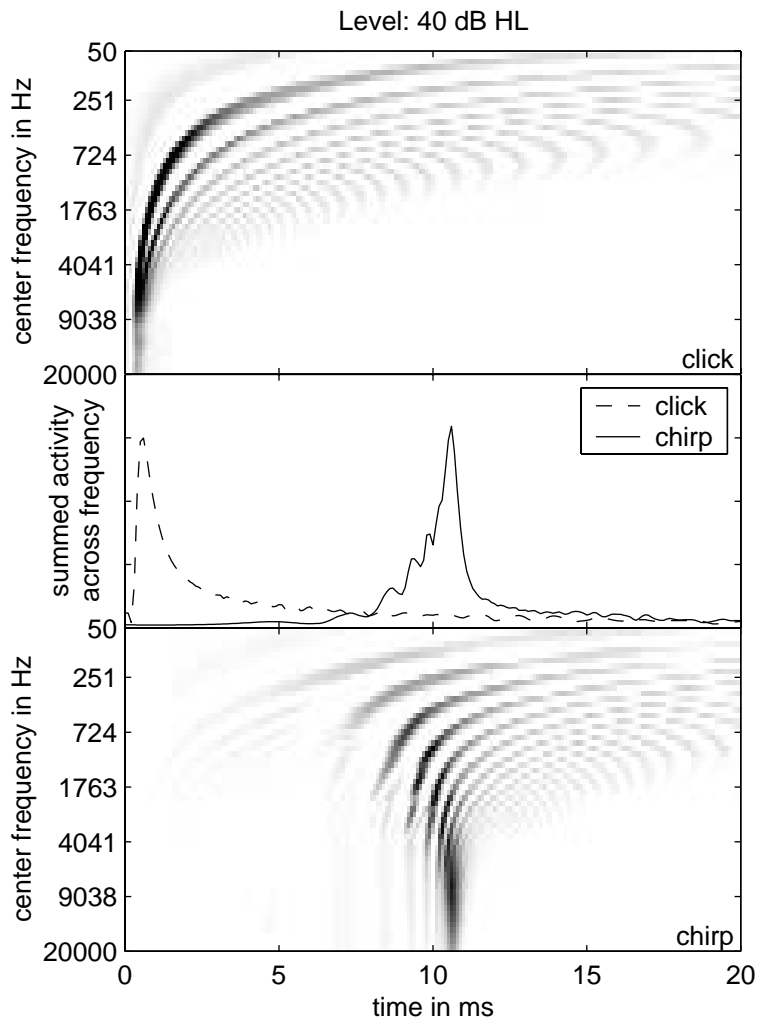
were presented in combination with masking noise. First, since the chirp stimulus has been designed to minimize phase canceling effects of evoked activity, the chirp responses in all high-pass noise masking conditions are of larger amplitude than the click responses where some of the activity is still phase canceled. The finding that the chirp causes a significantly larger response even at the lowest cut-off frequency of 0.5 kHz makes the stimulus interesting for clinical use in assessing low-frequency information. Second, it was observed that for the chirp, wave-V amplitude increases with increasing noise cut-off frequency while for the click it tends to saturate for frequency regions above about 1 kHz. This further confirms that, for the chirp, neural activity from the entire frequency range contributes to the evoked response



**Figure 3.11:** Mean wave-V amplitudes (left panel) and latencies (right panel) for the conditions from Fig. 3.10. Error bars indicate standard deviation.

while, for the click, neural activity is less synchronized across frequencies because of the travel-time differences on the basilar membrane.

To illustrate the effects of peripheral processing associated with cochlear response times, Fig. 3.12 shows simulated neural activity patterns for click stimulation (top panel) and for broadband chirp stimulation (bottom panel). After middle ear filtering, approximated by a second-order band-pass filter with cut-off frequencies of 0.3 and 7 kHz, the stimuli served as input to the recently developed auditory-nerve (AN) model by Heinz *et al.* (2001). The model is a modification of the physiologically based AN model by Zhang *et al.* (2001) which was developed for the cat and is itself an extension of the original Carney (1993) model. The model by Heinz *et al.* (2001) uses a human cochlear map according to Greenwood (1990), and the auditory filter bandwidth has been matched to humans based on psychophysical estimates of auditory filters (Glasberg and Moore, 1990). This model includes effects of level-dependent tuning, level-dependent phase, compression, suppression and fast nonlinear dynamics on the response. The model is specifically designed to describe the time-varying discharge rate of the AN fibers for a given characteristic frequency (CF). A set of 60 model CFs were used in the present study. The CFs ranged from 0.1 to 10 kHz and were spaced according to a human cochlear map (Greenwood, 1990). Details of the processing stages can be found in Heinz *et al.* (2001).



**Figure 3.12:** Neural activity patterns for click (top panel) and chirp stimulations (bottom panel), obtained with the AN model of [Heinz et al. \(2001\)](#). The middle panel represents the summed neural activity across the frequency channels.

A stimulus level of 40 dB HL was used for the following simulation since this level was also used in the experiments. It can be seen that for the click (top panel) the maxima of the neural activity are delayed at the lower frequencies relative to those at higher frequencies, while for the chirp (bottom panel) the maxima of neural excitation are more synchronized across frequency. Note that the simulation is based on a functional basilar-membrane model assuming Gammatone filters that treat the cochlea essentially as a bank of bandpass filters and ignore cochlear hydrodynamics while the chirp was developed based on spatial variations in cochlear geometry and mechanics that underlie the frequency-position map. Nevertheless, the resulting delay-line characteristic obtained with the different modeling concepts is similar. Within the framework of the [Heinz et al. \(2001\)](#) model, the latencies directly result from the filter bandwidths assumed. The middle panel of Fig. 3.12 shows the summed activity for the click (dashed curve) and for the chirp (solid curve), calculated by simply adding up the



activities of all 60 frequency channels. The peak activity for the chirp is shifted (by about the duration of the chirp) relative to that of the click, and has a larger amplitude. This agrees qualitatively with experimentally obtained compound action potentials (CAPs) for click and chirp stimulation in the guinea pig (Shore and Nuttall, 1985).

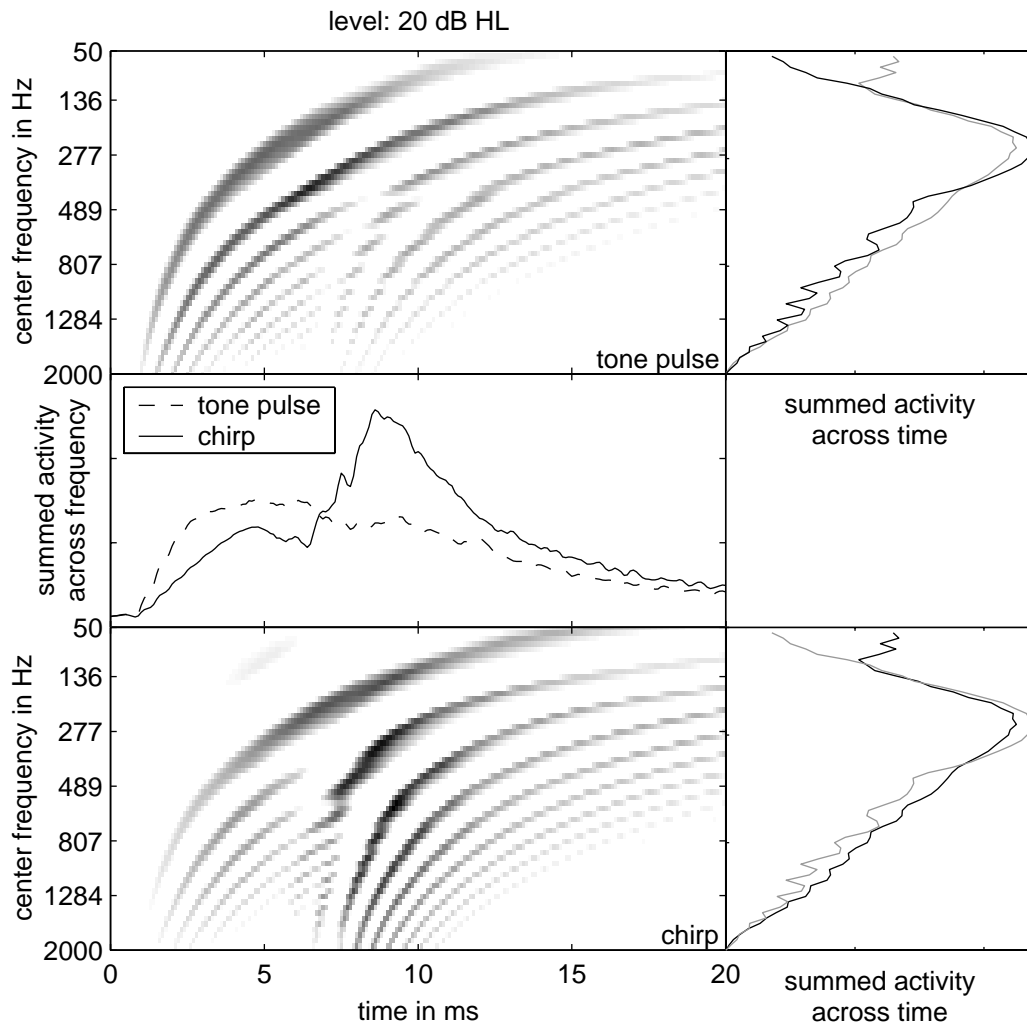
As was shown in the left panel of Fig. 3.7, if contributions from one-octave wide frequency regions are investigated, there is no significant difference in efficiency anymore between chirp and click stimulation. Derived responses were found to be of essentially the same amplitude for the two stimuli. Similar results were obtained with the notched-noise masker (right panel of Fig. 3.7). Apparently, the better synchronization of neural activity obtained with the chirp is *not* sufficient to produce an advantage if only one-octave wide frequency regions (or less) contribute to the response.

However, in all conditions tested in the present study, the broadband chirp was found to be equally effective as or better than the traditional click stimulus. In the conditions of high-pass masking alone the chirp always elicited a higher response than the click. The chirp may therefore be considered as the stimulus of choice for retrieving frequency-specific information as long as a broadband stimulus in combination with masking noise is considered.

### 3.4.2 Low-frequency chirp versus tone pulse

The results obtained with the narrow-band stimuli showed that a slight change in the phase characteristic can lead to large differences in the evoked ABR. Both the low-frequency chirp and the tone pulse consisted of only three half-waves and no ramps were used to introduce a smooth on- and offset. Thus, these stimuli exhibit a broadened magnitude spectrum compared to, e.g., tones of longer duration. It is important to note that the goal of the present study was not to develop an optimal *narrow-band* stimulus for retrieving low-frequency information from ABR. It is still controversial what narrow-band stimulus might represent the best choice. Instead, the present study attempts to emphasize the role of cochlear processing for the formation of ABR, particularly at low frequencies.

To illustrate the differences between the two narrow-band stimuli at the level of auditory-nerve processing, Fig. 3.13 shows the corresponding neural activity patterns obtained with the model described above. The top and bottom panels show the simulated output activity for tone-pulse and narrow-band chirp, respectively. A stimulus level of 20 dB HL was used in this case. The simulations nicely demonstrate the main characteristics observed in the



**Figure 3.13:** Neural activity patterns for the 250-Hz tone pulse (top panel) and the low-frequency chirp (bottom panel), obtained with the same model as in Fig. 3.12. The middle panel represents the synchronized neural activity integrated across the frequency channels. The two panels on the right show corresponding neural activity within each frequency channel, integrated over time, in order to estimate the neural “excitation pattern” obtained with the tone pulse (top) and the chirp (bottom). For direct comparison, each of these panels also shows the pattern for the other stimulus, as indicated by the Gray curve.

data. First, for the chirp neural activity is more dominated near the end of the stimulus while it is shifted towards earlier times for the tone pulse. Second, the activity is temporally more concentrated and higher in amplitude for the chirp than for the tone pulse. Both observations are summarized in the corresponding neural activities summed

across frequency as shown in the middle panel of Fig. 3.13. In order to also illustrate the amount of neural excitation as a function of frequency, the two panels on the right show the corresponding activity, integrated over time, in all excited frequency channels. For direct comparison, each of these panels also shows the pattern for the other stimulus, as indicated by the gray curve. The “excitation patterns” for the two stimuli are very similar. As for the magnitude spectra (from Fig. 3.2) the largest differences occur at frequencies below about 100 Hz where the chirp has more energy than the tone pulse. However, as can be seen in Fig. 3.13, this energy does only contribute to the synchronized activity at times *after* the maximum of synchronization. The simulations therefore suggest that the degree of synchronization but not the higher amount of overall neural activity is mainly responsible for the larger evoked response amplitude of the chirp.

However, for a deeper understanding of the relationship between peripheral processing and the formation of ABR patterns, additional assumptions about the neural processing at brainstem level as well as assumptions about the shape of the “unitary” responses of the different contributing neurons need to be considered (de Boer, 1975; Elberling, 1976; Goldstein and Kiang, 1958). Such modeling was beyond the scope of the present paper but is currently investigated (Dau, 2001).

### 3.4.3 Limitations of the chirp

There are some limitations of using the chirp that are considered in the following. For any given individual subject, the chirp designed from published functions regarding distance, frequency and temporal maps in the cochlea is not necessarily optimal for that individual. That is, there probably is significant variation from individual to individual in the cochlear response time between frequency regions. Thus, the chirp may represent a compensation that is optimized for some mean delay of a group of individuals. Amplitude differences between individuals or between cochlear regions within a given individual may be reflecting how well the chirp represents the true cochlear response times across and within individuals and not solely the amount of activation. This issue may become more problematic when impaired cochlea are assessed where cochlear filter characteristics might vary as a function of the degree of damage. Thus, if the chirp response amplitude is being used to assess the amount of neural activity, those differences may also be reflecting the differences in the appropriateness of the chirp used.

However, in a series of accompanying experiments we found that a variation of the chirp parameters (such as the sweeping rate) that may result from the variability of physical parameters (such as the stiffness constant of the basilar membrane) had only little effect on the evoked response amplitude of the same subject. In addition, we also generated chirps based on estimates of the delay line characteristic from transiently evoked otoacoustic emissions (Shera and Guinan, 2000). In contrast, our original chirp was generated on the basis of parameters that were derived from post mortem experiments. Again, even though these two chirps differed considerably in duration and sweeping rate, the corresponding evoked potential amplitude hardly differed from one another for any subject tested so far. This suggests that the evoked potential measure does not seem sensitive enough to reflect *slight* changes in the physical properties of the cochlea. It seems sufficient to roughly match the (inverse) delay line characteristic of the cochlea in order to produce a much better response than the click (as long as the effective frequency range of the contributing neurons is not too narrow).

Of course, since chirps make an assumption of the response times along the cochlea, the actual delays down the cochlea in terms of peak activity can not be estimated *directly* because it has already been established by the parameters of the chirp. In contrast, such information is directly available when using the click-evoked derived bands because one observes to a large extent the “natural” response of the system and differences due to cochlear damage can be studied. However, in the case of the chirp, such peak activity can be derived indirectly as the difference from the assumed characteristic in the chirp parameters. There is no information loss when using the chirp instead of the click.

#### 3.4.4 Relation to other methods of frequency specificity

Since the sweeping rate of the chirp is determined by basilar-membrane traveling-wave properties, the chosen “nominal” edge frequencies determine the duration of the chirp. A chirp consisting of only three halfwaves (starting and ending with zero crossings) will always be stretched over about two octaves, irrespective of the frequency region. It is possible that ABR obtained with more frequency selective stimuli allow for a better estimate of frequency specific information. For example, tone pulses as suggested by Davis (1976) with rise and fall times equal to two cycles of the stimulus frequency and a plateau time equal to one cycle may represent a good compromise between frequency specificity and sufficient syn-

chronization capability of the stimulus. A direct comparison between such a stimulus and a narrow-band chirp was not attempted in the present study.

However, as mentioned in the introduction, ABR elicited by stimulation with brief tone pulses of frequencies below 2 kHz has been shown to be only a poor predictor of low-frequency behavioral threshold, at least if presented without additional high-pass noise masking. Also, responses to longer-duration low-frequency tones, known as frequency following responses (FFR) do not represent frequency-specific information but most likely reflect synchronized activity from basally located regions (Janssen *et al.*, 1991; Dau, 2001) and can only be detected at relatively high stimulation levels above about 70 dB SPL (e.g., Hou and Lipscomb, 1979; Batra *et al.*, 1986). It has been suggested that instead of investigating transiently evoked responses and FFR, steady-state responses such as amplitude modulation following responses (AMFR) may serve as a better estimate of frequency specific information even though the interpretation of the steady-state responses may be complicated because activity from brainstem and cortical generators are superimposed (e.g., Kuwada *et al.*, 1986; Griffiths and Chambers, 1991).

Whatever stimulus might represent the best choice for assessing frequency specific information from evoked potentials, particularly at low frequencies, the results from the present study suggest that the chirp stimulus produces a higher neural synchrony than other stimuli of similar magnitude spectrum. This was demonstrated for the broadband chirp in comparison with the click as well as for the band-limited chirp in comparison with the tone pulse.

### 3.5 Summary and conclusions

- The broadband chirp elicited a larger wave-V amplitude than the click in the unmasked condition as well as in all masking conditions where high-pass noise was presented in addition to the signal. The results demonstrated that the increased synchrony obtained with the chirp stretches over the entire frequency region. The chirp may be particularly interesting for clinical use in the low-frequency region below about 0.5–1 kHz.
- The derived responses obtained with high-pass noise masking as well as the responses using notched-noise maskers indicate that the gain in synchrony within frequency regions

of about one octave is not sufficient for the chirp to produce a significantly larger response amplitude than the click.

- The low-frequency chirp elicited a larger wave-V amplitude at low and medium levels than a tone pulse with similar duration and magnitude spectrum. Wave-V latency differed by about 5 ms for these two stimuli. These observations as well as the differences between click and broadband chirp could be qualitatively explained in terms of the simulated neural activity patterns in the auditory periphery using a computational AN model.
- Overall, the results further demonstrate the importance of cochlear processing for the formation of ABR. In order to obtain a deeper understanding of these effects, modeling work is needed reflecting signal processing at cochlear level and at subsequent brainstem stages, as well as assumptions about the contributions of single unit activity at these stages to the far-field response.

## Chapter 4

# Searching for the optimal stimulus eliciting auditory brainstem responses in humans

## Abstract

This study examines auditory brainstem responses (ABR) elicited by rising frequency chirps. Two new chirp stimuli were developed and designed such as to compensate for basilar-membrane (BM) group-delay differences across frequency, in order to maximize neural synchrony at cochlear level. One chirp, referred to as the O-chirp in the present study, was based on estimates of human BM group delays derived from stimulus-frequency otoacoustic (SFOAE) at a level of 40 dB SPL [Shera and Guinan, in *Recent Developments in Auditory Mechanics*, (2000)]. The other chirp, referred to here as the A-chirp, was derived from latency functions fitted to tone-burst-evoked ABR wave V data over a wide range of stimulus levels and frequencies [Neely *et al.*, " J. Acoust. Soc. Am. **83**(2), 652–656 (1988)]. In this case, a set of level-dependent chirps, was generated. The chirp-evoked responses, particularly wave-V amplitude and latency, were compared with click responses and with responses obtained with the "original" chirp as defined in Dau *et al.* [J. Acoust. Soc. Am. **107**(3), 1530–1540], referred to as the M-chirp, which was developed on the basis of a (linear) cochlea model. The main hypothesis was that, at low and medium stimulation levels, the new chirps might produce a larger response than the original one whose parameters were essentially based on high-level BM data. The main results of the present study are as follows: (i) all chirps evoked a higher wave-V amplitude than the click. (ii) Surprisingly little differences occurred between the O-chirp and the original M-chirp for low and medium levels, indicating that SFOAE may only provide a relatively rough estimate of BM group delay. (iii) The A-chirp produced the largest responses, particularly at the lowest stimulation levels. This chirp might therefore be particularly interesting for clinical applications.

## 4.1 Introduction

Transient stimuli like clicks are commonly used in electrophysiological research of the human auditory system to elicit synchronized auditory brainstem responses (ABR). The click is the most common stimulus used in recording the ABR, whether for neurodiagnostic or audiologic purposes. However, in the cochlea, the response to a click is not entirely synchronous: the peak of the response occurs several milliseconds later in low-frequency channels than it does in high-frequency channels (von Békésy, 1960). The reason for this is that, as a result of



the change of stiffness along the cochlear partition, the phase velocity of the traveling wave depends in a characteristic way upon location, which causes spatial dispersion. It takes more time for the low-frequency region to reach maximal displacement at the apical end of the cochlea. As a consequence, electrophysiological responses to broadband transients like clicks appear to be largely generated by the synchronized activity of the high-frequency channels on their own.

[Don and Eggermont \(1978\)](#) measured human ABR in response to clicks masked by high-pass noise with different cut-off frequencies. This masking technique revealed that the latencies in response to low-frequency stimuli are delayed relative to high frequencies. [Don and Eggermont](#) concluded that there must be contributions to the ABR from all regions of the cochlea, although the response is dominated by contributions from the 2–3 octaves towards the basal end.

[Dau et al. \(2000\)](#) and [Wegner and Dau \(2002\)](#) recently demonstrated that upward chirps can affect wave V in the human ABR. Their chirp was designed to produce simultaneous displacement maxima along the cochlear partition by compensating for frequency-dependent traveling-time differences. Their equations determining the temporal course of the chirp were derived on the basis of a linear cochlea model ([de Boer, 1980](#)) and were calculated to be the inverse of the delay-line characteristic of the human cochlear partition. The fundamental relationship between stimulus frequency and place of maximum displacement was taken from [Greenwood \(1990\)](#). ABR evoked by the broadband chirp showed a larger wave-V amplitude than corresponding click-evoked responses. [Dau et al. \(2000\)](#) demonstrated that the ABR is not an electrophysiological event purely evoked by onset or offset of an acoustic stimulus, but that an appropriate temporal organization, determined by basilar-membrane (BM) traveling-wave properties, can significantly increase synchrony of neural discharges. [Wegner and Dau \(2002\)](#) demonstrated that such a chirp is also very useful for retrieving frequency-specific information, particularly at low frequencies.

The model of [de Boer \(1980\)](#) is based upon the experimental observations of [von Békésy \(1960\)](#). Von Békésy's measurements were performed with the aid of a microscope to detect and measure the movements of cochlear structures. The movements had to be much larger than under the influence of "natural" sounds. In fact, [von Békésy](#) used very high sound levels, of the order of 120 to 140 dB SPL. In later studies, cochlear vibration patterns were measured with more sensitive techniques and under more natural circumstances. These

physiological studies of cochlea mechanics have established that the response of the BM to tones at characteristic frequency (CF) is generally nonlinear and compressive (e.g., Rhode, 1971; Sellick *et al.*, 1982; Robles *et al.*, 1986; Ruggero, 1992) at medium sound levels while it responds linearly to tones with a frequency well below CF (Sellick *et al.*, 1982). As a consequence of the compressive nonlinearity of the BM vibration pattern in the normal system, the traveling wave shows a sharply-tuned peak. In a damaged cochlea, the sharply-tuned tip of the tuning curve disappears, the threshold rises, and the intensity function becomes linear. In such a situation, this pattern of vibration is similar to the “insensitive” response, originally found by von Békésy.

Recently, Shera and Guinan (2000) developed an objective, noninvasive method for determining the frequency selectivity of cochlea tuning at low and moderate sound levels. The method is based on the measurement of stimulus-frequency otoacoustic emissions. Evoked otoacoustic emissions (OAEs) are sounds, recordable in the ear canal with low-noise microphones, that originate within the cochlea (Kemp, 1978). OAEs can be evoked with various stimuli, but the easiest to interpret are those evoked by a pure tone – stimulus-frequency OAEs (SFOAEs), so-called because they occur at the frequency of stimulation. According to the theory of coherent reflection filtering (Shera and Zweig, 1993; Zweig and Shera, 1995), SFOAE characteristics are directly and quantitatively related to the mechanical response of the inner ear, via coherent reflection from “random” impedance perturbations such as those of cochlear anatomy. Coherent reflection filtering predicts that SFOAE group delay  $\tau_{\text{SFOAE}}$  (defined as the negative slope of the SFOAE-phase versus frequency function) is determined by the group delay  $\tau_{\text{BM}}$  of the BM mechanical transfer function at its peak (evaluated at the cochlea location with CF equal to the stimulus frequency). Specifically, the theory implies that  $\tau_{\text{SFOAE}} = 2 \cdot \tau_{\text{BM}}$ . It is assumed that  $\tau_{\text{BM}}$  is related to cochlea-filter bandwidth since, at low levels, BM transfer functions manifest many of the characteristics of minimum-phase-shift filters (Zweig, 1976). In particular, their bandwidth and phase slopes are reciprocally related, with smaller bandwidths corresponding to steeper phase slopes (i.e., longer delays  $\tau_{\text{BM}}$ ). Estimates of cochlear frequency selectivity at low levels, obtained with the method of SFOAEs, will certainly differ from those values assumed in de Boer’s model. As a consequence, a corresponding chirp stimulus that, theoretically, compensates for frequency-dependent traveling-time differences, can be expected to differ in its waveform from the chirp developed by Dau *et al.* (2000).

Another, more straight-forward, approach to compensate for delays across frequency would be to base the chirp parameters on wave-V latency values obtained in tone burst evoked ABR data. [Gorga \*et al.\* \(1988\)](#) measured tone-burst-evoked ABRs over a wide range of stimulus levels and frequencies. Their wave-V-latency data were described by [Neely \*et al.\* \(1988\)](#) with an exponentially decreasing function:

$$\tau_b = a + bc^{-i}f^{-d}, \quad (4.1)$$

where  $i$  represents tone-burst intensity,  $f$  represents tone-burst frequency and  $a$ ,  $b$ ,  $c$ , and  $d$  are constants (cf. their Eqn. 1). They assumed that the first term, parameter  $a$ , represents the frequency and level independent neural component of the latency while the second term in Eqn. 4.1 reflects the mechanical component of the latency due to the propagation in the cochlea and thus representing BM group delay. By comparing the ABR data from [Gorga \*et al.\* \(1988\)](#) with tone-burst OAE data from [Norton and Neely \(1987\)](#) – who used a subgroup of the subjects from [Gorga \*et al.\* \(1988\)](#) – [Neely \*et al.\* \(1988\)](#) found a much larger inter- and intraindividual variability in the OAE data than in the ABR data suggesting that BM group delay might be better estimated with ABR than with OAE.

The current study deals with the development and test of new chirp stimuli in an attempt to find the optimal stimulus eliciting ABR in humans. Chirps are generated based on the recent SFOAE data by [Shera and Guinan](#) and on the ABR wave-V latency data by [Neely \*et al.\* \(1988\)](#). Chirp evoked responses are compared with results obtained with the “original” chirp by [Dau \*et al.\* \(2000\)](#) and with conventional click data, at various levels. The underlying hypothesis is that, at low stimulation levels, the new chirps might produce a better synchronization than the original chirp by [Dau \*et al.\* \(2000\)](#) since this latter chirp was derived on the basis of high-level BM data. Such results would be valuable for clinical applications using chirp-evoked ABR as an objective indicator of hearing threshold.

## 4.2 The chirp stimuli

### 4.2.1 OAE-based chirp stimulus (O-chirp)

The first new chirp stimulus is based on the experimental SFOAE data of [Shera and Guinan \(2000\)](#). They made experiments for stimulus frequencies in the range from 0.5 to 10 kHz in humans, at a level of 40 dB SPL. Emission group delays,  $\tau_{\text{SFOAE}}$ , were calculated and related

to BM group delays,  $\tau_{\text{BM}} = 0.5 \cdot \tau_{\text{SFOAE}}$ , as a function of CF. The data can be roughly approximated by the following function (Shera, personal communication):

$$\tau_{\text{BM, Shera}}(f) := t(f) = \frac{c}{\sqrt{f}}, \quad (4.2)$$

with the constant  $c = 0.15 \text{ Hz}^{-0.5}$ .  $\tau_{\text{BM, Shera}}$  can also be interpreted as reflecting the propagation time  $t(f)$  needed to arrive at the place of resonance for frequency  $f$ . In order to compensate for the dispersion across frequency, we introduced the variable transformation  $t \rightarrow t_0 - t$ , with  $t_0 = t(50 \text{ Hz})$ <sup>1</sup>, and calculated the inverse function  $f_{\text{O}}(t) = t^{-1}(f)$ :

$$f_{\text{O}}(t) = \left( \frac{c}{t_0 - t} \right)^2. \quad (4.3)$$

This function reflecting the change of the instantaneous frequency with time was then integrated over time to derive the instantaneous phase  $\varphi_{\text{O}}(t)$  of the resulting chirp:

$$\varphi_{\text{O}}(t) = 2\pi \int_0^t f_{\text{O}}(t') dt' \quad (4.4)$$

$$= 2\pi c^2 \left( \frac{1}{t_0 - t} - \frac{1}{t_0} \right). \quad (4.5)$$

The chirp stimulus is then given by:

$$s_{\text{O}}(t) = A_{\text{O}}(t) \sin(\varphi_{\text{O}}(t) - \varphi_{\text{O}}(t_0)), \quad (4.6)$$

whereby the time-dependent amplitude factor  $A_{\text{O}}(t)$  was chosen as:

$$A_{\text{O}}(t) = \sqrt{\frac{df_{\text{O}}(t)}{dt}} = \sqrt{\frac{2c^2}{(t_0 - t)^3}}, \quad (4.7)$$

in order to produce a stimulus with a flat magnitude spectrum. Since the stimulus  $s_{\text{O}}(t)$  is based on OAE data, it is referred to as the ‘‘O-chirp’’ throughout the present paper.

## 4.2.2 ABR based chirp stimulus (A-chirp)

The second chirp stimulus developed in this study, is based on the tone-burst-evoked ABR data from [Gorga \*et al.\* \(1988\)](#). They used tone bursts at ten frequencies (0.25, 0.5, 0.75,

<sup>1</sup> $t_0$  can be chosen somewhat arbitrarily. [Dau \*et al.\* \(2000\)](#) used  $t_0 = t(0 \text{ Hz})$ . Since this value is not defined within Eqn. 4.2, we used  $t_0 = t_{\text{BM, Shera}}(50 \text{ Hz})$ .

1, 1.5, 2, 3, 4, 6, and 8 kHz) and nine intensities (20 to 100 dB SPL in 10-dB steps). Their data were described by Neely *et al.* (1988) with an exponentially decreasing function (cf. Eqn. 4.1). Neely *et al.* (1988) assumed that the total wave-V latency represents the sum of mechanical and neural components. Since they assumed the neural component to be independent of frequency and level, they claimed the second term in Eqn. 4.1 to represent the component of ABR latency due to mechanical propagation within the cochlea. Thus they described the BM group delay by

$$\tau_{\text{BM,Neely}}(i, f) = bc^{-i}f^{-d}, \quad (4.8)$$

where  $i$  represents the tone-burst intensity (in dB SPL divided by 100),  $f$  represents tone-burst frequency (divided by 1 kHz), and  $b$ ,  $c$ , and  $d$  are constants with the values  $b = 12.9$  ms,  $c = 5.0$ , and  $d = 0.413$ , according to the data fit from Neely *et al.* (1988). It should be noted that this description of BM group delay depends on both frequency *and* level of the stimulus. As for the O-chirp, from this function the inverse function and the function for the instantaneous phase  $\varphi_{\text{A}}(i, t)$  can be calculated:<sup>2</sup>

$$\varphi_{\text{A}}(i, t) = \frac{2\pi (b \cdot c^{-i})^{\frac{1}{d}}}{\frac{1}{d} - 1} \left[ \frac{1}{(t_0(i) - t)^{\frac{1}{d}-1}} - \frac{1}{t_0(i)^{\frac{1}{d}-1}} \right] \quad (4.9)$$

The chirp stimulus is then given by:

$$s_{\text{A}}(i, t) = A_{\text{A}}(i, t) \sin(\varphi_{\text{A}}(i, t) - \varphi_{\text{A}}(i, t_0(i))), \quad (4.10)$$

whereby the time *and* intensity dependent amplitude factor  $A_{\text{A}}(i, t)$  was chosen as:

$$A_{\text{A}}(i, t) = \sqrt{\frac{(b \cdot c^{-i})^{\frac{1}{d}}}{d \cdot (t_0(i) - t)^{\frac{1}{d}+1}}} \quad (4.11)$$

in order to produce a stimulus with a flat magnitude spectrum. Throughout the present paper, the stimulus  $s_{\text{A}}(i, t)$  is referred to as the ‘‘A-chirp’’, since it is based on ABR data.

### 4.2.3 Comparison of the different chirp stimuli

The two new chirps are compared with the original one as defined in Dau *et al.* (2000). This original chirp was based on de Boer’s (1980) linear cochlea model. For direct comparison,

---

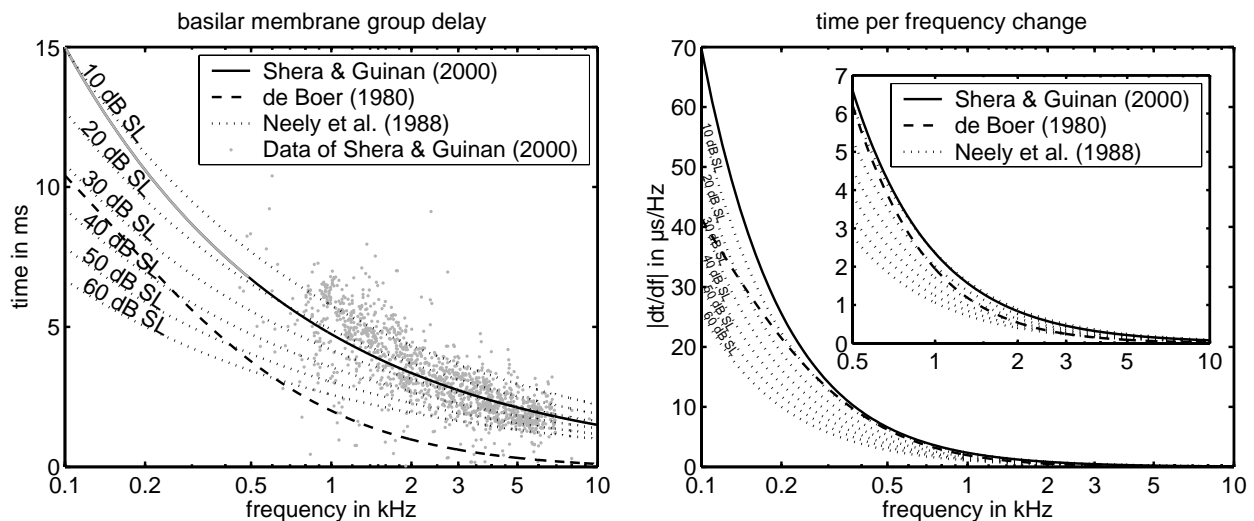
<sup>2</sup> $t_0(i)$  can again be chosen somewhat arbitrarily, but in this case it depends on the stimulus intensity  $i$ . We chose  $t_0(i) = \tau_{\text{BM,Neely}}(i, 50 \text{ Hz})$ .

the realization of the chirp with a flat magnitude spectrum is used, also defined in [Dau et al. \(2000\)](#). Since this chirp is based on a model, it is called “M-chirp” in the following.

Within [de Boer’s \(1980\)](#) model, the propagation time,  $t_{\text{deBoer}}(f)$ , needed to arrive at the place of resonance for the frequency  $f$ , is approximately given as:

$$t_{\text{deBoer}}(f) \propto (f + 165.4 \text{ Hz})^{-1.1}, \quad (4.12)$$

which clearly differs from Eqns. 4.2 and 4.8 representing the corresponding functions for the other chirps. The left panel of Fig. 4.1 illustrates the calculated BM group delays on the basis of Eqns. 4.2, 4.8 and 4.12. The group delays derived from the SFOAE experiments (solid curve) are about 2 – 5 ms larger than those predicted on the basis of [de Boer’s](#) model (dashed curve). The shaded dots represent the original BM group delay estimates of the SFOAE data of [Shera and Guinan \(2000\)](#). Such a shift towards higher values is reasonable,



**Figure 4.1:** Left: basilar-membrane group delay as a function of frequency. The shaded dots represent the original BM group delay estimates of [Shera and Guinan \(2000\)](#), derived from SFOAE data. The black solid line represents their data fit. This function is interpolated to lower frequencies (gray solid line). The dashed line represents the group delay on the basis of the linear cochlea model by [de Boer \(1980\)](#). The dotted lines represent the group delays predicted from [Neely et al. \(1988\)](#) for stimulus intensities from 10 dB SL (top dotted curve) up to 60 dB SL (bottom dotted curve). Right: time per frequency within the different chirps. This was directly calculated from the data in the left panel. The second axis is a replot of the frequency range from 0.5 to 10 kHz with a rescaled ordinate.

at least qualitatively, since frequency selectivity is increased at lower levels and this should be accompanied by a larger group delay at all frequencies. The group delay estimates predicted by the equation of Neely *et al.* (1988) depend on intensity and vary by about 1 ms at 10 kHz up to about 8 ms at 0.1 kHz (dotted curves).

For the generation of the chirps, not the absolute values but the *change* of group delay with frequency is important. Thus, the derivative of the group-delay versus frequency function is shown in the left panel of Fig. 4.1. The main differences between the OAE-based curve and the model-based prediction occur at low frequencies, but differences are also present at frequencies above 500 Hz. This means that the instantaneous frequency of the corresponding O-chirp will vary more slowly than that of the M-chirp, particularly at low frequencies. The corresponding ABR-based curves clearly differ from the other curves. There is a large variation with stimulus level. For example, at higher levels, the resulting A-chirp will vary much faster than in the case of the O- and the M-chirp, resulting in much shorter chirp durations, as shown in Fig. 4.2 and described further below.

## 4.3 Method

### 4.3.1 Subjects

Nine normal-hearing subjects (two female and seven male) with no history of hearing problems and audiometric thresholds of 15 dB HL or better participated in the experiments. All subjects were between 28 and 38 years of age, and either volunteered or were paid for the experiment.

### 4.3.2 Apparatus

The experiments were carried out with a PC-based computer system which controlled stimulus presentation and recording of evoked potentials. A DSP-card (Ariel DSP32C) converted the digitally generated stimulus (25 kHz, 16 bit) to an analogous waveform. The output of the DSP card was connected to a digitally controlled audiometric amplifier, which presented the stimulus through an insert earphone (Etymotic Research ER-2) to the subject.

Electroencephalic activity was recorded from the scalp via silver/silver chloride electrodes, attached to the vertex (positive) and the ipsilateral mastoid (negative). The fore-

head served as the site for the ground electrode. Inter-electrode impedance was maintained below 5 k $\Omega$ . Responses were amplified (80 dB) and filtered (30–3000 Hz) with a commercially available EEG preamplifier (TDT DB4/HS4). The amplified signal was digitized by the DSP-card (25 kHz, 16 bit), which also performed artifact rejection and signal averaging. Responses were recorded for 37 ms following the stimulus onset. Off-line filtering (digital low-pass, 1600 Hz, order 4) was done to suppress noise.

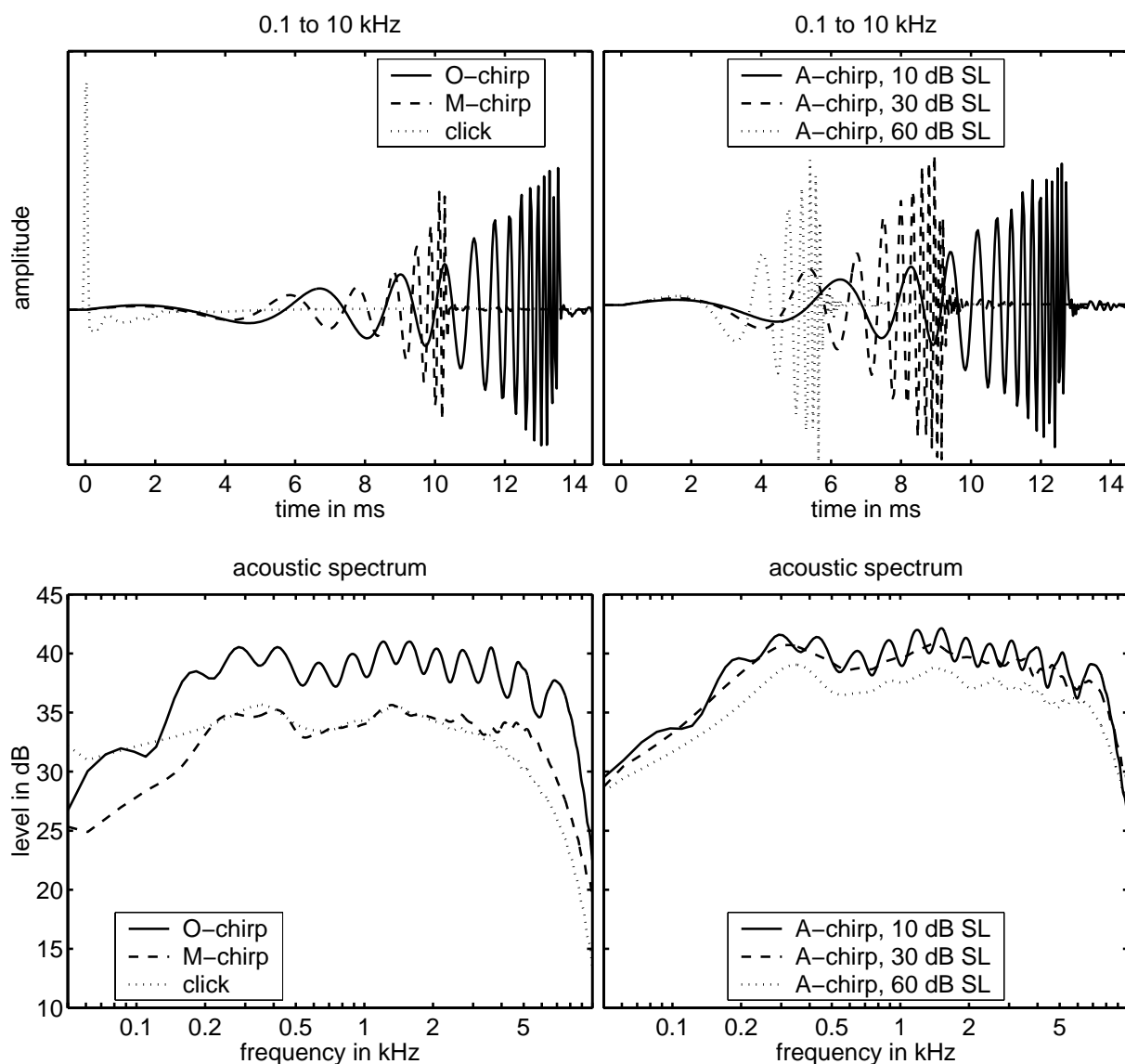
### 4.3.3 Stimuli and procedure

Chirps as described in Sec. 4.2 were used as stimuli. The nominal edge frequencies of the chirps were 0.1 and 10 kHz resulting in durations of 13.52 ms for the O-chirp and 10.32 ms for the M-chirp. The durations for the A-chirp varied between 5.72 ms for 60 dB SL and 12.72 ms for 10 dB SL. To compare results with standard ABR measurements, an 80- $\mu$ s click stimulus was generated. The upper left panel of Fig. 4.2 shows the acoustic waveforms of the O-chirp, the M-chirp, and the click stimulus. The waveforms of the A-chirp for stimulation levels of 10, 30, and 60 dB SL are depicted in the upper right panel. The corresponding acoustic spectra are given in the lower panels. They were obtained by coupling the ER-2 insert earphone to a Brüel and Kjær ear simulator (type 4157) with a 1/2-in. condenser microphone (type 4134), a 2669 preamplifier, and a 2610 measuring amplifier. The spectra were derived from fast Fourier transforms (FFTs) of 100-trial time-domain averages of the stimulus over an analysis frame of 2048 samples using a sampling rate of 25 kHz. The waveforms were not windowed prior to FFT.

Since [Shera and Guinan \(2000\)](#) collected only very few data points below 0.5 kHz, Eqn. 4.2 may represent only a poor description of the behavior in this frequency region. Therefore, additional O- and M-chirp stimuli were generated with nominal edge frequencies of 0.5 and 10 kHz. The corresponding durations were 5.24 ms for the O-chirp and 3.68 ms for the M-chirp. Figure 4.3 shows the acoustic waveforms (left panel) and spectra (right panel) of these stimuli. Fig. 4.3.

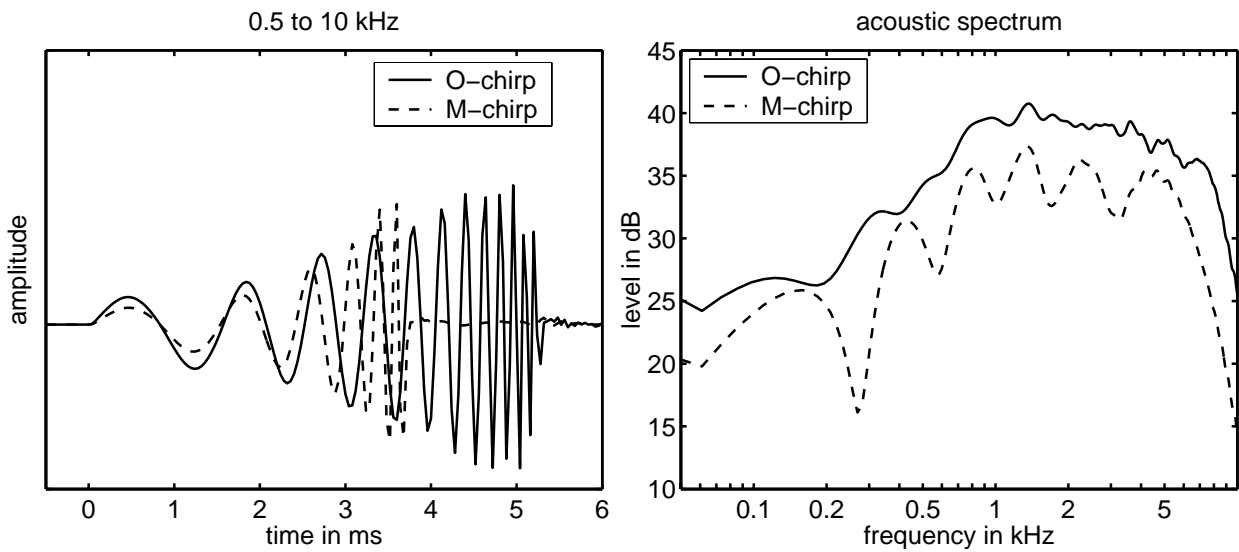
For all stimuli, the presentation level was varied between 10 and 60 dB SL in 10-dB steps. To determine the sensation level for the different stimuli, the absolute hearing thresholds were measured individually with an adaptive three interval three alternative forced choice (one up, two down) procedure.





**Figure 4.2:** Temporal waveforms (upper panel) and corresponding acoustic spectra (lower panel) of the stimuli ranging from 0.1–10 kHz. The left panels show the O-chirp, the M-chirp, and the click stimulus, which are indicated as solid, dashed, and dotted functions, respectively. The right panels show the corresponding functions for the level dependent A-chirps generated for 10, 30, and 60 dB SL. Different levels were indicated by different line styles. For better comparison all waveforms were measured at a level of 100 dB peSPL.

The subject lay on a couch in an electrically shielded, soundproof room, and electrodes were attached. The subject was instructed to keep movement at a minimum, and to sleep if possible. The lights were turned off at the beginning of the session. Each session lasted



**Figure 4.3:** Temporal waveforms (left panel) and corresponding acoustic spectra (right panel) of the M-chirp and O-chirp with nominal edge frequencies of 0.5 and 10 kHz. O-chirp and M-chirp are indicated by solid and dashed lines, respectively.

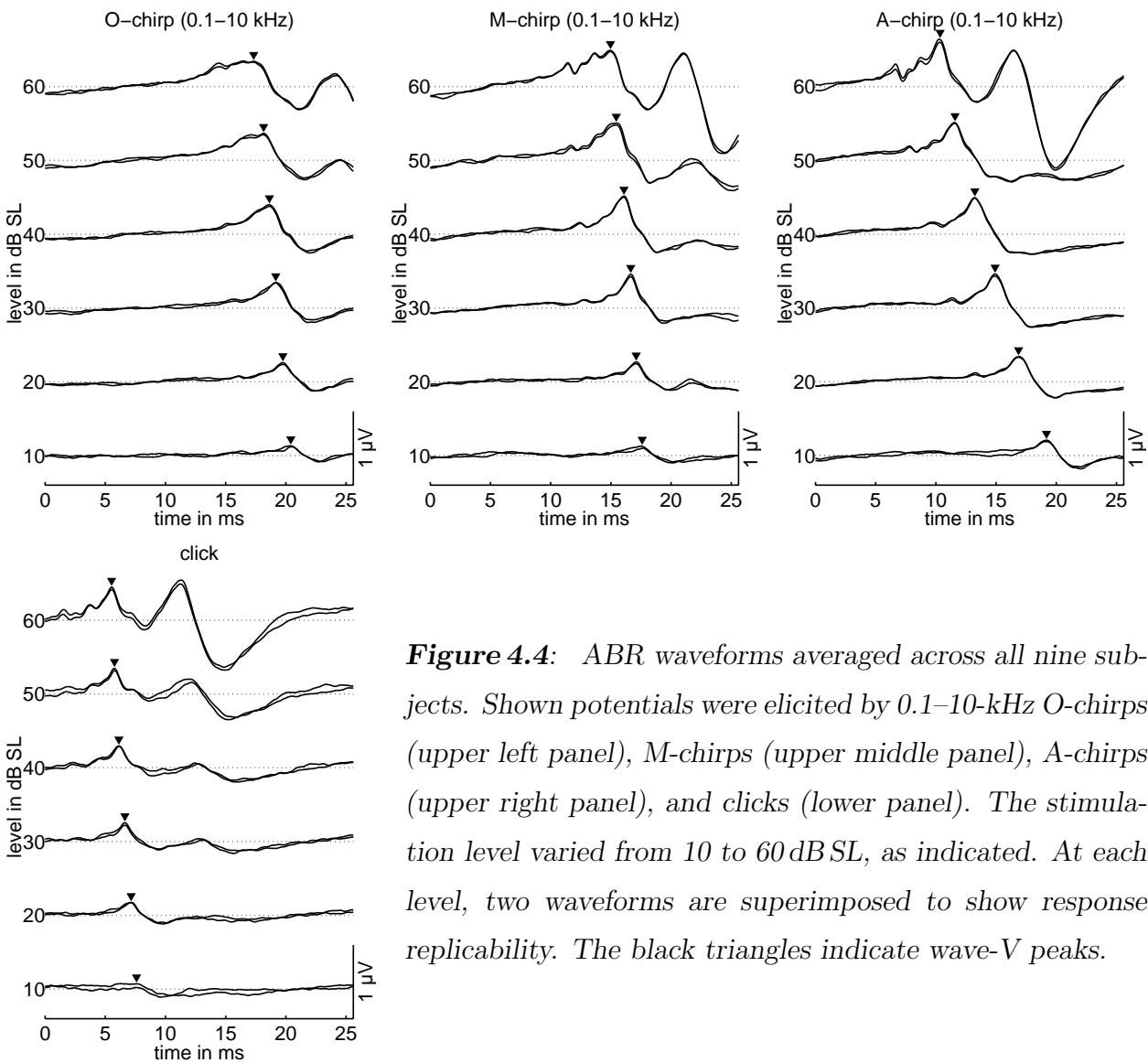
between one and two hours, depending on the subject's ability to remain still. The ear of stimulation was chosen randomly, i.e., for each subject one ear was chosen and then maintained. The acoustic signals were delivered at a repetition rate of 20 Hz for all stimulus conditions. A temporal jitter of  $\pm 2$  ms was introduced to minimize response superimposition from preceding stimuli. Thus the time interval between the onsets of two successive stimuli varied randomly and equally distributed between 48 and 62 ms. Each trial consisted of 3000 averages. For each stimulus condition, two independent trials were stored in separate buffers. These are illustrated as superimposed waveforms in the figures to show response replicability.

#### 4.3.4 Statistical analysis

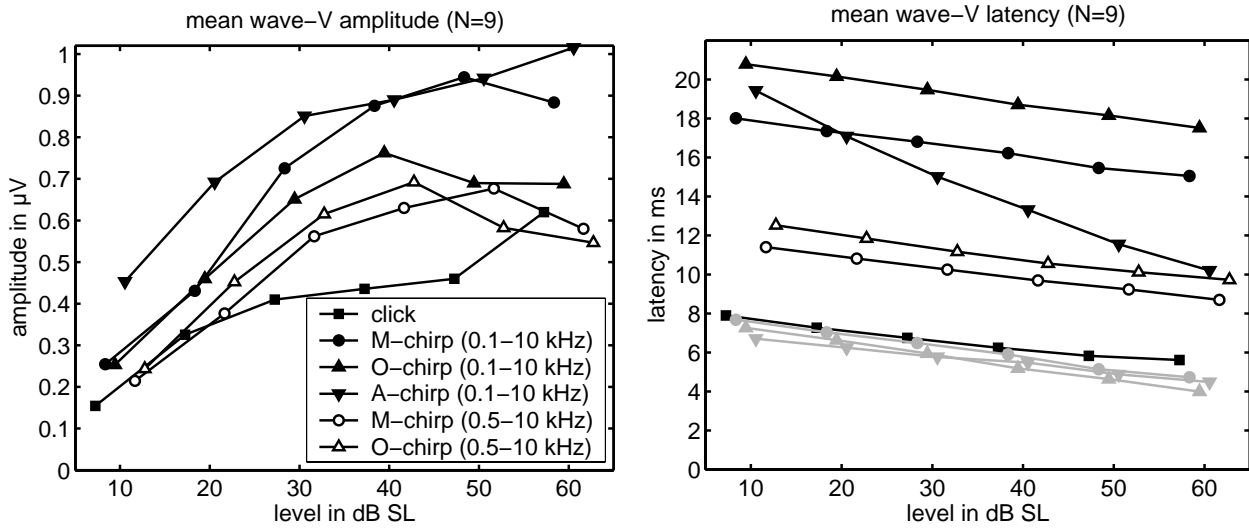
Wave-V peak-to-peak amplitude was analyzed in all stimulus conditions. The amplitude was measured from the peak to the largest negativity following it. For each condition, wave-V amplitude was averaged across subjects. A Wilcoxon matched-pairs signed-rank test ( $\alpha = 0.05$ ) was performed to test whether the response amplitude differed significantly for two comparison stimuli.

## 4.4 Results

Figure 4.4 shows mean ABR, averaged across all nine subjects, obtained with the OAE-based (0.1-10-kHz) O-chirp (upper left panel), the original model-based M-chirp (upper middle panel), the ABR-based A-chirp (upper right panel), and the click (lower panel). Results for different stimulus levels are indicated on separate axes along the ordinate, and labeled with the corresponding sensation level (dB SL). Wave-V peaks are marked by small black triangles. Wave V is the only peak that can be observed in *all* stimulus conditions. For the O-chirp, no earlier waves are present, even not at the highest stimulation levels. In contrast,



**Figure 4.4:** ABR waveforms averaged across all nine subjects. Shown potentials were elicited by 0.1–10-kHz O-chirps (upper left panel), M-chirps (upper middle panel), A-chirps (upper right panel), and clicks (lower panel). The stimulation level varied from 10 to 60 dB SL, as indicated. At each level, two waveforms are superimposed to show response replicability. The black triangles indicate wave-V peaks.



**Figure 4.5:** Average ABR data for wave-V amplitude (left panel) and latency (right panel), as a function of stimulation level. Different symbols indicate different stimulus conditions. ■: click; ●: M-chirp starting with 100 Hz; ▲: O-chirp starting with 100 Hz; ▼: A-chirp starting with 100 Hz; ○: M-chirp starting with 500 Hz; △: O-chirp starting with 500 Hz. The shaded symbols in the right panel indicate the offset latencies for the corresponding stimuli. For better visibility, the symbols are slightly shifted along the abscissa.

for the M-chirp, A-chirp, and the click, waves I and III become visible at the highest levels. Interestingly, for the A-chirp, wave I is even visible down to a level of 20 dB SL.

Fig. 4.5 (left panel) summarizes the quantitative values for mean wave-V amplitude as a function of the stimulation level. First of all, the click-evoked wave-V amplitude, represented by the filled squares, is always smaller than that obtained with any of the broadband chirps, represented by the other filled symbols. For example, the M-chirp (filled circles), leads to amplitude values that are more than twice the values for the click at most stimulus levels. This agrees well with the results found in [Dau et al. \(2000\)](#). At the lowest stimulation level, the A-chirp (filled downward triangles) even evoked an amplitude that is about three times as large as that for the click. The amplitude-level function for the A-chirp looks like shifted by about  $0.44 \mu\text{V}$  with respect to the click curve. For the A- and M-chirp, statistical analysis revealed significantly larger amplitudes than for the click at *all* stimulus levels while for the (broadband) O-chirp the difference was significant only for 10 and 40 dB SL.

Now consider the results for the O-chirp (filled upward triangles) in comparison to the original chirp (M-chirp) having in mind that the O-chirp was based on 40 dB SPL OAE data while the M-chirp was based on a (linear) cochlea model fitted to high-level data. At levels of 40 dB SL and above, wave-V amplitude is smaller for the O-chirp than for the M-chirp, while at the lower levels, wave-V amplitude is about the same for the two stimuli. Statistical analysis of the amplitude data revealed significant differences between the O- and M-chirp only for levels of 50 and 60 dB SL, where wave-V amplitude for the M-chirp is higher. The results for the smaller chirp bandwidth ranging from 0.5-10 kHz are indicated by the corresponding open symbols. Results for the O- and the M-chirp are given by triangles and circles, respectively. The response waveforms are not shown explicitly for these two conditions. Statistical analysis revealed a significant difference between the two chirps only for a level of 20 dB SL, where the O-chirp evoked a higher wave-V amplitude than does the M-chirp.

Next consider the results for the A-chirp (filled downward triangles) in comparison to the original one. The A-chirp revealed a larger wave V amplitude than the M-chirp (and any other stimulus tested here) at nearly all stimulation levels. However, statistical analysis revealed significant differences between A- and M-chirp only for low stimulation levels (10 and 20 dB SL).

Comparison of the results for the A-chirp with the one for the O-chirp shows that the A-chirp revealed a higher wave-V amplitude than the O-chirp at all stimulation levels used here. In this case statistical analysis results in significant differences for low and high stimulation levels (10, 20, 50, and 60 dB SL).

The right panel of Fig. 4.5 shows the mean wave-V latency behavior obtained with the different stimuli. Except for the A-chirp, all functions are roughly in parallel to each other but shifted relative to each other by some amount. For these functions, the latency decreases by about 2–3 ms for a 50-dB level change (from 60 to 10 dB SL), which is consistent with literature data (e.g., [Hoth and Lenarz, 1994](#)). The main differences between the functions correspond to the differences in the respective stimulus durations, as is illustrated by the shaded functions in the same panel of the figure. They indicate the latency values for the three broadband chirps relative to stimulus *offset* instead of stimulus onset. The very similar values in this view are consistent with the idea behind the chirp paradigm that, ideally, the displacement maxima on the BM should occur in *all* channels at the same time, and thus,

the latencies for the chirp and the click should be similar if expressed relative to stimulus offset. Thus, since the duration of the A-chirp changes strongly with level, this must be directly reflected in a relatively steep function if expressed relative to stimulus onset.

## 4.5 Discussion

In previous studies it was demonstrated that an upward chirp can evoke a significantly larger wave-V amplitude than the conventional click (Dau *et al.*, 2000; Wegner and Dau, 2002). The equations defining this upward chirp, called the M-chirp in the present study, were calculated to be the inverse of the delay-line characteristic of the cochlear partition on the basis of the linear cochlea model by de Boer (1980). However, this model does not take cochlear nonlinearities into account which are mainly responsible for the sharpening of the filters at low and medium levels. Therefore, the model presumably underestimates real BM group delays and it could be assumed that the M-chirp may not represent the optimal choice at low and medium stimulation levels.

The intention of the present study was to design and test new chirp stimuli that were optimized for lower stimulation levels and that might potentially cause an even larger neural synchronization than the M-chirp, at least at the low stimulation levels. Since ABR are often used as an objective indicator of hearing threshold, such an optimized stimulus might be very interesting for clinical applications. Two different strategies for the generation of the new stimuli were used: the first one was based on OAE data by Shera and colleagues (Shera and Guinan, 2000; Shera *et al.*, 2002), recorded at a stimulation level of 40 dB SPL in humans. Their derived estimates for the BM group delays from SFOAE group delays were used in the present study for the generation of a new chirp, the O-chirp. A function for the frequency dependence was approximated to the group-delay data, and the chirp was defined such that it compensates for the BM group-delay differences across frequency. The course of the sweeping rate and the resulting temporal waveform of the O-chirp differed considerably from those for the M-chirp. Our assumption was that, at low levels, the O-chirp should produce a larger wave-V response amplitude than the M-chirp (and the click), while, at high levels, it should be less effective than the M-chirp.

However, the experimental results of the present study showed surprisingly little differences in response amplitude between the O- and the M-chirp. For the frequency region above

500 Hz, where reliable SFOAE data were available, the two (0.5–10-kHz) chirps produced about the same wave-V amplitude. No significant advantage could be obtained with the O-chirp at the lower levels (except for 20 dB SL), and also no advantage could be observed for the M-chirp at the high levels. The results were slightly different for the broadband conditions (0.1–10 kHz). Here the M-chirp produced a higher potential amplitude than the corresponding broadband O-chirp at the two highest levels. This indicates that, for the O-chirp, our assumed extrapolated group-delay-versus-frequency function at the very low frequencies (0.1–0.5 kHz) probably does not match the real system very well. At these very low frequencies, the function underlying the M-chirp seems to represent a better choice. In principle, the experimental data confirm the observations from our earlier studies that an appropriate stimulation at frequencies below 500 Hz can have a significant influence on the overall neural synchrony. What might be the reason(s) for the similar results obtained with the O- and the M-chirp? It is not clear whether OAEs really represent a good estimate of cochlear group delay. The SFOAE group-delay data in [Shera and Guinan \(2000\)](#) show a very large variability (cf. Fig. 4.1, left panel). Earlier studies also found large variations of OAE data within and across subjects, especially when compared with ABR data ([Neely et al., 1988](#)). Even if assuming that individual OAE data would result in good estimates of individual BM group delay, for any given individual subject, the chirp designed from (approximated) functions in the cochlea is probably not optimal for that individual.

The second strategy used in the present study was based on ABR measurements. The idea was to use tone-burst evoked ABR wave-V latencies measured at different frequencies and levels. [Neely et al. \(1988\)](#) fitted a (frequency and level dependent) exponential function to the data set of [Gorga et al. \(1988\)](#). The corresponding chirp derived in the present study, the A-chirp, was defined such as to compensate for the delay differences across frequency. The resulting duration for the A-chirp varied much with stimulation level while the duration of O- and the M-chirp were level independent. Since this chirp was “optimized” for *all* levels, the hypothesis was that the A-chirp should be advantageous over the M- and the O-chirp at all levels. Indeed, wave-V amplitudes for the A-chirp were found to be higher than for any of the other stimuli of the present study, at any stimulation level. The disadvantage of this approach is that it is not possible to discriminate between mechanical and neural components: wave-V latency always represents the sum of these two components and so does the fit provided by [Neely et al. \(1988\)](#). Thus it may be possible that there is a level

dependent neural component involved, which is in contrast to the assumptions made in [Neely \*et al.\* \(1988\)](#) where all level dependency was attributed to the mechanical component while the neural component was assumed to be constant. The advantage of the second approach is that it may represent the optimal stimulation when attempting to achieve maximal neural synchronization at brainstem level, at least in the normal-hearing system. The advantage is particularly large at very low stimulation levels. The A-chirp may therefore be very interesting and valuable for clinical applications using chirp-evoked ABR as an objective indicator of hearing threshold.

## 4.6 Summary and conclusions

- Two new chirp stimuli compensating for BM spatial dispersion were developed. One chirp (O-chirp) was based on BM group-delay estimates obtained with SFOAE ([Shera and Guinan, 2000](#)). The other one (A-chirp) was based on BM group-delay estimates from level dependent tone-pulse evoked ABR wave-V latencies ([Gorga \*et al.\*, 1988](#); [Neely \*et al.\*, 1988](#)). ABR obtained with these chirps were compared with click responses and with responses from the original chirp stimulus (M-chirp) developed in [Dau \*et al.\* \(2000\)](#) which is based on the [de Boer \(1980\)](#)'s linear cochlea model.
- Mean ABR data showed essentially no significant differences between O-chirp and M-chirp wave-V amplitude at most stimulation levels. However, comparing A-chirp and M-chirp, mean ABR data showed larger potentials for the A-chirp, whereby differences were significant only at low levels.
- Since all chirps caused a larger response amplitude than the click, they are interesting for clinical application since they include activity from lower-frequency regions which do not effectively contribute in the case of the click.
- Overall, the A-chirp may be suggested as reflecting the stimulus of choice for future studies since it leads – even if not always significantly – to the largest response amplitudes in all conditions.



## Chapter 5

# Modeling auditory evoked middle latency responses (MLR)

## Abstract

A model for the generation of middle latency responses (MLR) is presented. It is an extension of the auditory brainstem response (ABR) model introduced by [Dau](#) [J. Acoust. Soc. Am. (2003)]. The model uses the concept that evoked potentials can be described by convolution of an elementary unit waveform (unitary response) with the instantaneous discharge rate function for the corresponding unit. The instantaneous discharge rate functions are calculated with the auditory nerve (AN) model developed by [Heinz \*et al.\*](#) [ARLO **2**(3), 91–96 (2001)]. The summed excitation across frequency is convolved with an unitary response function, which is assumed to reflect contributions from different cell populations along the auditory pathway. The unitary response is derived by deconvolving high-level click-evoked MLR with the corresponding summed AN activity. Predicted potential patterns are compared with corresponding experimental data for clicks and chirps at different stimulus levels and for clicks at a wide range of different repetition rates. The results show that, despite the large simplifications made, some of the main characteristics of the MLR and steady state responses (SSR) can be accounted for reasonably well. However, the strong response amplitude for repetition rates around 40 Hz cannot fully be accounted for by the above linear system's approach, indicating that additional processes need to be involved that are not considered in the current version of the model.

## 5.1 Introduction

Evoked responses represent the summation of activity from many neurons, recorded from electrodes placed on the surface of the head (e.g., [Jewett, 1970](#)), i.e., remote to the individual neurons. Auditory evoked potentials can be recorded from all levels of the auditory system. They are usually grouped by the time of occurrence after the onset of the stimulus, and this grouping corresponds roughly to the site of generation. The click-evoked middle-latency response (MLR) generally consists of various prominent components. Two prominent peaks are the N19 (or N<sub>a</sub>) and the P30 (or P<sub>a</sub>), which occur approximately 19 and 30 ms after stimulus onset. Source analysis of EEG and MEG data, as well as intracranial studies, indicate that the N19–P30 complex originates from the medial portion of Heschl's gyrus in the primary auditory cortex ([Scherg \*et al.\*, 1989](#); [Scherg and von Cramon, 1990](#); [Liègeois-](#)

Chauvel *et al.*, 1994; Gutschalk *et al.*, 1999; Rupp *et al.*, 2002), but which cells exactly generate the various components of the MLR is not yet fully understood.

Goldstein and Kiang (1958) introduced the concept that remote responses generated by auditory-nerve (AN) neurons can theoretically be described by the convolution of an elementary unit waveform, the unitary response (UR), with the instantaneous rate function at which the AN cell discharges in response to the stimulus. Following a more general description of this concept, Melcher and Kiang (1996) suggested that the potential produced by *any* cell in the auditory pathway can be described by the convolution of the instantaneous rate function with the unitary potential which is produced each time this cell discharges.

Recently, Dau (2003) used this concept to model auditory brainstem responses (ABR) and frequency following responses (FFR) in humans. He used a nonlinear computational auditory-nerve model (Heinz *et al.*, 2001) to calculate the instantaneous discharge rate of AN fibers over a wide frequency range. The summed activity across frequency was convolved with a unitary response which is assumed to reflect contributions from different cell populations within the auditory brainstem. His unitary response function was calculated by deconvolution of mean experimental click ABR data, taken from (Dau *et al.*, 2000), with the summed neural activity pattern for the click. He predicted potential patterns for a number of stimulus and level conditions and compared the predictions with experimental data. He found the main characteristics and key observations from the data to be reflected in the simulations.

In the present study, the modeling approach by Dau (2003) is extended to human middle-latency responses. The new model is tested with click and chirp stimuli for various level conditions. Also, for the click, the model is tested for a set of repetition rates in the range from 4 to 521 Hz in order to investigate the transition between transient and steady-state stimulation. The unitary response waveform is estimated via deconvolution for one specific stimulus condition, and then kept constant for all stimulus conditions. A similar approach was used by Rupp *et al.* (2002) in an MEG study. Clicks, up- and down-chirps were used to evoke middle-latency auditory evoked fields (MAEF): they deconvolved the different empirical source waveforms with spike probability functions simulated with another preprocessing model, the auditory image model (AIM; Patterson *et al.*, 1992, 1995). They found that the source waves for all stimulus conditions could be explained with the same unitary response

function. However, they only investigated one stimulation level and used one repetition rate in that study.

## 5.2 Method

### 5.2.1 Subjects and stimulation paradigm

Five normal-hearing subjects (one female and four male) with no history of hearing problems and audiometric thresholds of 15 dB HL or better participated in the experiments. All subjects were between 28 and 30 years of age, and either volunteered or were paid for the experiment. The subjects were tested while sitting in a comfortable reclining chair in a acoustically and electrically shielded room. They were instructed to watch videos, displayed on a LC display in front of them.

Stimuli were generated digitally (50 kHz, 16 bit) and downloaded to a DSP card (Ariel DSP32C) in the host computer. The DSP card converted the digital stimulus to an analogous waveform. The output of the DSP card was connected to a digitally controlled audiometric amplifier, which presented the stimulus monaurally through an insert earphone (Etymotic Research ER-2) to the subject.

For examining the level dependency of click- and chirp-evoked MLR, stimuli were either 0.1 ms clicks or flat-spectrum chirps (0.1–10 kHz, 10.32 ms) as defined in [Dau \*et al.\* \(2000\)](#). The stimulus onset asynchrony was 122.88 ms resulting in a stimulus repetition rate of about 8.1 Hz. Stimuli were presented with levels from 55 to 100 dB peSPL in 15-dB steps.

To examine click-rate effects, 0.1-ms clicks with a level of 100 dB SPL were used. The stimulus onset asynchronies varied between 1.92 ms and 245.76 ms. Exact values and the corresponding approximated repetition rates are given in [Table 5.1](#).

**Table 5.1:** *Stimulus onset asynchronies (SOA) and corresponding approximated repetition rates for the click stimuli used in this study.*

SOA in ms	1.92	3.84	7.68	15.36	24.80	30.72	61.44	122.88	245.76
Rate in Hz	520.8	260.4	130.2	65.1	40.3	32.6	16.3	8.1	4.1

### 5.2.2 Recording

Electroencephalic activity was recorded from the scalp via silver/silver chloride electrodes. Three active electrodes were placed at the left (A1) and the right (A2) mastoid as well as one centimeter below the inion (Iz). The common reference electrode was placed at the vertex (Cz). The forehead (Fpz) served as the site for the ground electrode. Electrode labels are according to the 10-20-system (Jasper, 1957). Inter-electrode impedance – measured at a test signal frequency of 30 Hz – was maintained well below 5 k $\Omega$ ; common values were 2–3 k $\Omega$ . A DC-coupled differential amplifier (Synamps 5803) was used to record the auditory evoked potentials (AEP). The headstage amplified the EEG signal by a factor 150. Further amplification was achieved by the main amplifier (factor 33), resulting in a total amplification of about 74 dB. During the data acquisition, an antialiasing-low-pass filter with a cut-off frequency of 3 kHz was used before digitizing (50 kHz, 16 bit). Filtering, artifact analysis and averaging were done off-line. The recording interval consisted of 6144 samples in the time interval from 0 to 122.8 ms relative to stimulus onset.<sup>1</sup> For each stimulus condition, 10000 epochs were averaged. For each subject the data acquisition was done in three different sessions. In each session, five to six runs were presented in random order, resulting in a total duration of about two hours.

### 5.2.3 Data analysis

Before averaging all sweeps across all subjects, the single sweeps were filtered with a linear phase FIR bandpass filter with 200 taps (Granzow *et al.*, 2001). An iterated weighted average of the filtered sweeps was computed across all subjects for all stimulus conditions (Riedel *et al.*, 2001).

For the level dependent click- and chirp-evoked MLR, filter cut-off frequencies were chosen to 20 and 300 Hz. Amplitude differences between ABR wave V and MLR N19, and between MLR N19 and P30 were analyzed for each stimulus condition. Latencies were measured for wave V, N19 and P30.

The click-rate dependent MLR were bandpass filtered (2–1500 Hz) and transformed to the frequency domain. The absolute value of the FFT bin representing the click rate was

---

<sup>1</sup>For a click rate of 4 Hz the recording interval comprised 12288 samples in the time interval from 0 to 245.6 ms

calculated and converted to microvolts, giving the amplitude value for the different click rates.

## 5.3 The model for MLR generation

### 5.3.1 The general modeling approach

Auditory evoked potentials (AEP) can be assumed to represent the sum of potentials  $v_i$  produced by individual cells  $i$  in response to the stimulus  $s$ , combined across all corresponding cells along the auditory pathway (Goldstein and Kiang, 1958; Melcher and Kiang, 1996):

$$\text{AEP}(t, \bar{x}_1, \bar{x}_2, s) = \sum_i v_i(t, \bar{x}_1, \bar{x}_2, s) \quad (5.1)$$

The potential  $v_i$  depends on time ( $t$ ), the locations of the recording electrodes ( $\bar{x}_1, \bar{x}_2$ ), and the stimulus ( $s$ ).  $v_i$  produced by any given cell can be separated into two terms: the first one is the instantaneous discharge rate function,  $r_i(t, s)$ , at which the cell discharges in response to the stimulus. The second one is the potential produced between  $\bar{x}_1$  and  $\bar{x}_2$  each time the cell discharges, the unitary response  $u_i(t, \bar{x}_1, \bar{x}_2)$ , which is independent from the stimulus ( $s$ ). Thus,

$$v_i = r_i(t, s) * u_i(t, \bar{x}_1, \bar{x}_2) \quad (5.2)$$

where  $*$  denotes convolution. The unitary potential waveform depends on the morphological and electrical properties of the cell within the context of the entire head (Melcher and Kiang, 1996). Melcher and Kiang (1996) suggested considering groups of cells collectively, based on established anatomical and physiological criteria. The AEP can then be written as sum of potentials  $V_p$ , produced by any population  $p$  of cells:

$$\text{AEP}(t, \bar{x}_1, \bar{x}_2, s) = \sum_{p=1}^P V_p(t, \bar{x}_1, \bar{x}_2, s) \quad (5.3)$$

$$= \sum_{p=1}^P (R_p(t, s) * u_p(t, \bar{x}_1, \bar{x}_2)) \quad (5.4)$$

where  $P$  is the number of contributing populations.  $u_p$  is the unitary potential for each cell in the population, since it can be assumed that each cell in a given population produces the same unitary potential. This assumption is reasonable, since cells of a given physio-anatomical

type generally have, by definition, similar morphological and electrical properties.  $R_p$  is the sum over the discharge-rate functions,  $r_{i,p}$ , of the cells within the population:

$$R_p(t, s) = \sum_{i=1}^{N_p} r_{i,p}(t, s) \quad (5.5)$$

where  $N_p$  is the number of cells in population  $p$ . Thus, in order to calculate the AEP waveform for a particular stimulus and electrode configuration, one needs to know the unitary waveforms,  $u_p$ , of the cellular generators which generate the specific extrema in the AEP, as well as the corresponding summed discharge rate functions,  $R_p$ , for the cell populations  $p$ .

As in (Dau, 2003) for the ABR model, it was specifically assumed in the present study that the instantaneous discharge rate functions are the same as at the level of the auditory nerve, such that  $R_p(t, s) =: R_{AN}(t, s)$ . It follows from Eqn. 5.4 that:

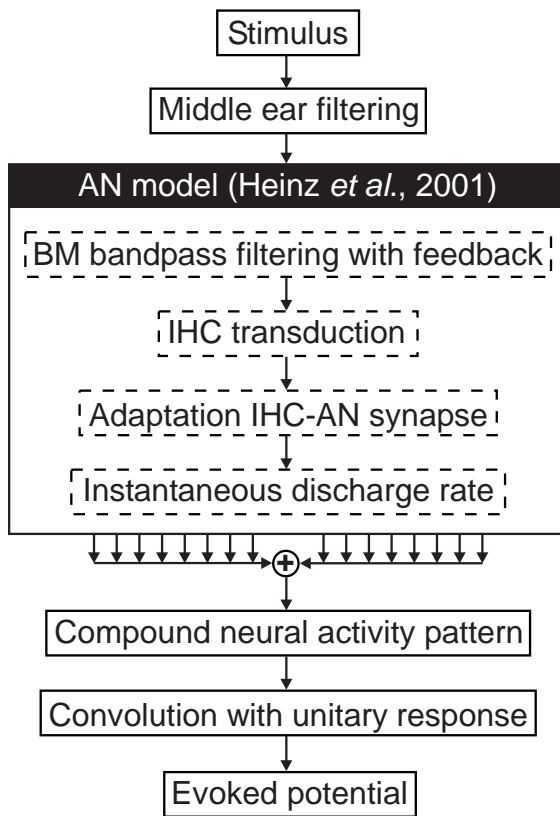
$$\text{AEP}(t, \bar{x}_1, \bar{x}_2, s) = R_{AN}(t, s) * \sum_{p=1}^P u_p(t, \bar{x}_1, \bar{x}_2) \quad (5.6)$$

Thus, the differences between the contributions from the different populations to the scalp potential are assumed to be reflected in the shape of the corresponding individual unitary responses  $u_p$ .

Figure 5.1 shows the block diagram of the overall model used in the present study for the simulation of AEP. The first stage roughly simulates the middle-ear transformation. The incoming stimulus is filtered by a second-order Butterworth bandpass filter with cut-off frequencies of 0.3 and 7 kHz, respectively. It follows the processing through the computational AN model described below. The next stage in the model calculates the summed neural activity pattern,  $R_{AN}$ , by adding up all discharge rate functions across characteristic frequency (CF). This pattern is then convolved with the unitary response waveform described below. The output of the model represents the simulated evoked response pattern for the considered stimulus.

### 5.3.2 The auditory-nerve model

The computational AN model developed by Heinz *et al.* (2001) was used – without variation – in the present study to calculate the instantaneous discharge rate functions. As indicated in Fig. 5.1, the model consists of nonlinear basilar membrane (BM) filtering, inner hair-cell (IHC) transduction, adaptation at the IHC-AN synapse and generation of the instantaneous



**Figure 5.1:** Structure of the model for the generation of auditory evoked responses. After middle-ear filtering, the stimuli are processed through the auditory-nerve model by Heinz et al. (2001). The instantaneous discharge rate functions are then summed across frequency. This summed activity pattern ( $R$ ) is convolved with the unitary response function ( $\sum_p u_p$ ) resulting in the modeled AEP. Details are described in the text.

discharge rate as a function of CF. A detailed description of the model and its implementation can be found in Heinz et al. (2001). In the following, some of the main characteristics are summarized.

The input stage is a filter bank that simulates the mechanical tuning of the BM. The model uses a human cochlear map according to Greenwood (1990), and the auditory filter bandwidths have been matched to humans based on psychophysical estimates of auditory filters (Glasberg and Moore, 1990). The parameters of these filters vary continuously as a function of stimulus level via a feedback mechanism, simulating the compressive nonlinearity associated with the mechanics of the BM. Level-dependent gain (compression), bandwidth, and phase properties are implemented with a control path that varies the gain and bandwidth of tuning in the signal-path filter. The properties required of the feedback signal are similar to the response properties of outer hair cells (OHC): a compressive magnitude response appears near the characteristic frequency, with the maximum compression occurring at CF and an essentially linear response far apart from CF. The compression starts at 20 dB SPL and is maximal at 40 dB SPL, consistent with physiological data (Ruggero et al., 1997). The amount of compression (or cochlear-amplifier gain) in the model is largest for high CFs and



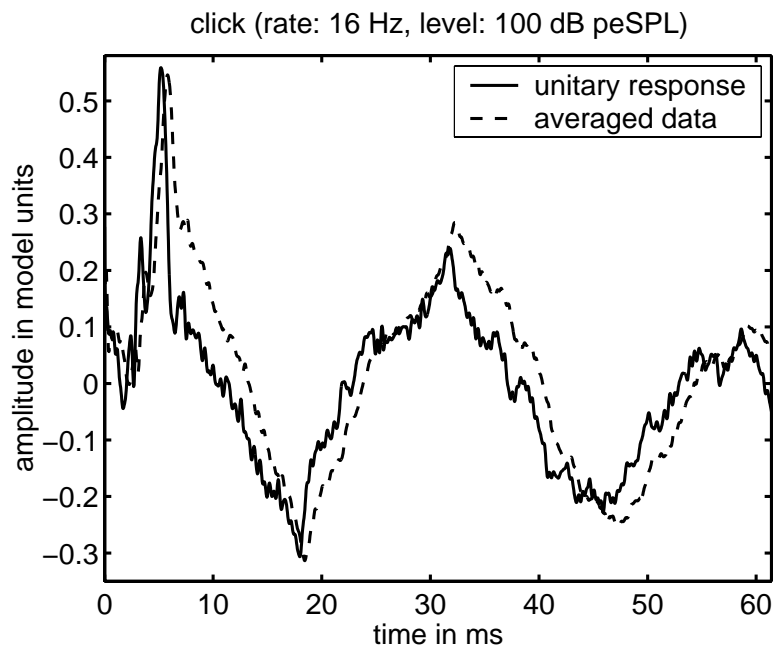
decreases towards lower frequencies, consistent with data from basal and apical turns of the chinchilla cochlea (Ruggero *et al.*, 1997; Cooper and Rhode, 1997).

The time-varying AN discharge rate is calculated by passing the output of the signal-path filter through an asymmetric saturating nonlinearity, a low-pass filter, and a synapse model. The saturating nonlinearity and the low-pass filter produce response properties associated with inner hair cell (IHC) transduction, whereas the synapse model includes adaptation effects such as the extended dynamic range at onset relative to the steady-state response.

For the simulations of the present study, a set of 500 model CFs was used. The CFs ranged from 0.1 to 10 kHz, and were spaced according to a human cochlear map (Greenwood, 1990).

### 5.3.3 The unitary response function

As described in Sec. 5.3.1, the stimulus dependent neural excitation function for the generation of MLR was assumed to be given by the single function  $R_{AN}(t, s)$ . Figure 5.2 (solid curve) shows the overall unitary response used in the present study. The functions were calculated by deconvolution of the mean experimental click MLR data (channel IZ, SOA 61.44 ms) at 100 dB peSPL (dashed curve) with the summed neural activity pattern for the click, generated by the AN model. Tikhonov regularization was applied (Tikhonov, 1963; Hansen, 1997) to achieve a stable and smooth solution for the inverse problem inherent in



**Figure 5.2:** The solid curve represents the unitary response function used in the present study. It was calculated by deconvolution of the mean experimental click data (dashed line) with the summed neural activity pattern for the click, generated by the AN model. In the present study, this unitary response function is used for all stimuli at all levels.

deconvolution. The extraction of appropriate and objective regularization parameters was based on the generalized cross-correlation function (GCV). All calculations were done in MATLAB R12.1 (The Mathworks, Inc.). The analysis tools for the regularization problems, including the GCV function to extract the optimal parameters, were provided by Hansen (1994). The average click data is indicated by the dashed line in Fig. 5.2. The data (dashed curve) show the typical pattern with clear waves V, N19 and P30 and P50 with latencies that correspond well to a large body of literature data. The obtained unitary response shows some more high-frequency components than the data. This can be observed at the peaks, which are narrower than those in the data, and in the fine structure of the UR. This is as expected since the convolution reflects a smoothing operation.

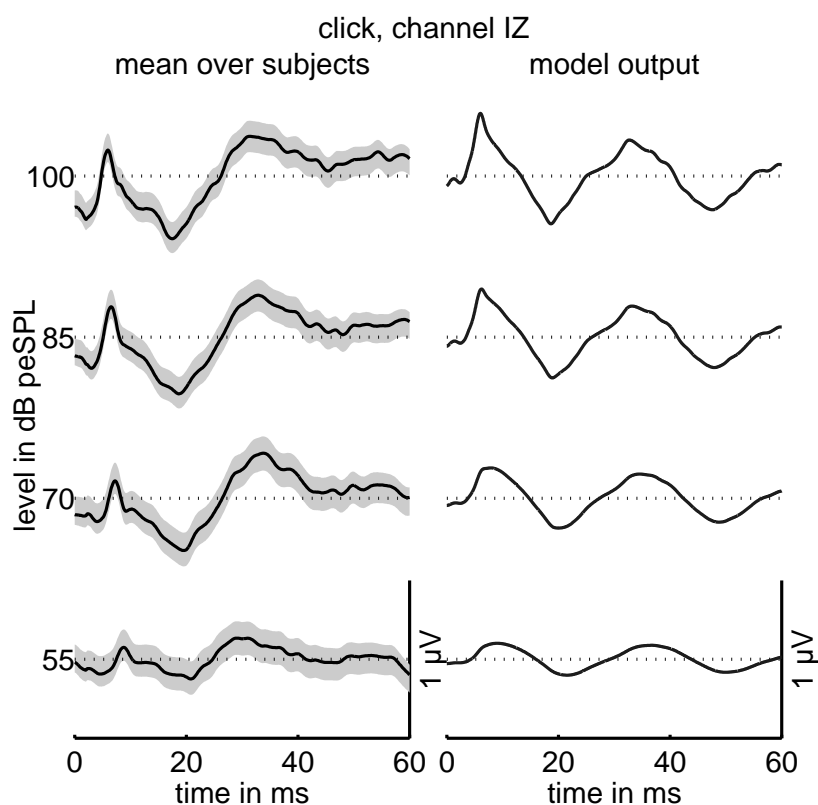
Within the present study, the above overall unitary response function is used for the different input stimuli and at any level, implying the assumption of linearity at this stage of processing. All nonlinearity in the model is restricted to the processing of the stimulus-dependent rate functions in the AN model. In the following, it is investigated whether the model accounts for the intensity and rate dependent aspects of the MLR data.

## 5.4 Results

### 5.4.1 Click- and chirp-evoked MLR as a function of level

Figure 5.3 shows the experimental click data (left panel) in comparison to the simulated patterns (right panel). Results for different stimulus levels are indicated on separate axes along the ordinate, and labeled with the corresponding stimulus level in dB peSPL. The gray area in the left panel indicates  $\pm 3$  standard errors. Simulations in the right panel were calculated with the unitary response shown in Fig. 5.2.

Wave V, N19 and P30 are the only peaks that can clearly be detected in all conditions. For both, experimental and simulated data, the amplitudes of the V–N19 and the N19–P30 complex decrease with decreasing level. For high stimulation levels, the general shape of the model output matches the one of the data quite well. This is not surprising, since the UR was calculated based on high-level click data. However, at lower levels the model fails to simulate the relatively sharp wave-V peak in the data; the simulated wave-V peak is

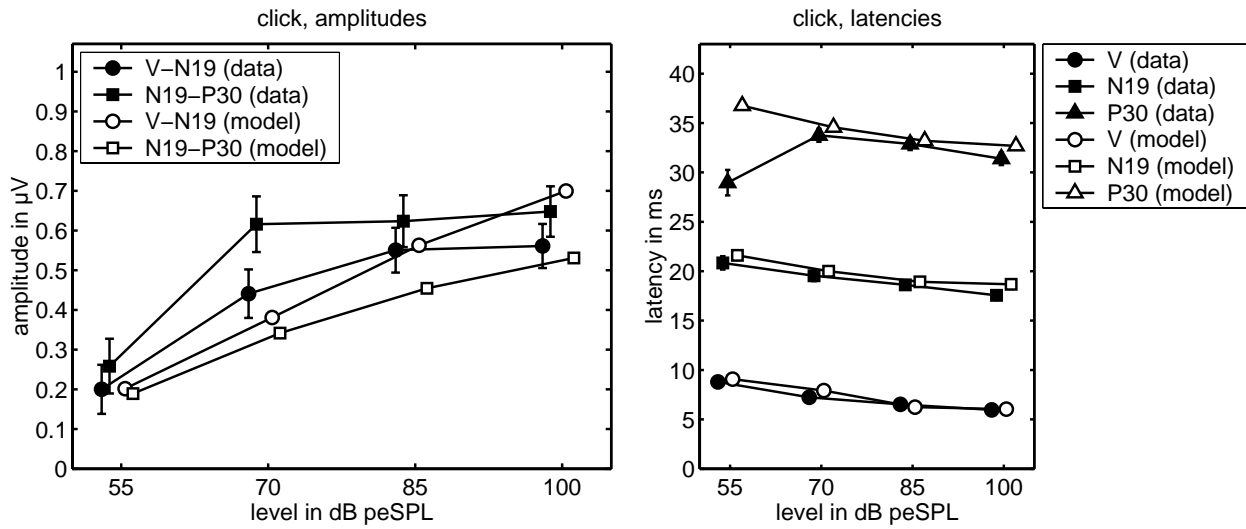


**Figure 5.3:** Left panel: click evoked MLR averaged across all five subjects (channel IZ, bandpass filtered with cut-off frequencies of 20 and 300 Hz). The stimulation level varied from 55 to 100 dB peSPL, as indicated. Stimulus presentation rate was about 8/s. The gray area indicates  $\pm 3$  standard errors. The right panel shows the corresponding model output.

somewhat smoother. The N19 and P30 components are modeled reasonably well at low levels.

Results from a more quantitative analysis are depicted in Fig. 5.4. Measured data are indicated by closed symbols, whereby open symbols represent model data. The left panel shows the amplitudes. Wave-V–N19 amplitudes (circles) are described well for the levels 55 and 85 dB peSPL. However, the saturation of the amplitude in the data at high levels is not reflected in the model. The measured amplitude of the N19–P30 complex (squares) also shows a saturation effect at high levels. Again, the model does not predict this saturation effect, and generally underestimates the amplitudes in the data. The right panel of Fig. 5.4 shows the latencies for wave V, N19 and P30. Except for the latency of the P30 peak at 55 dB peSPL, which is smaller by about 8 ms than for the model data, the model describes the latency data very well. However, it should be noted that the measured latencies for P30 at 55 dB peSPL do not agree well with typical literature data (e.g., [Picton \*et al.\*, 1974](#)).

In Fig. 5.5, the experimental chirp data (left panel) are compared to the corresponding simulated patterns (right panel). The parameters were as in Fig. 5.3. The agreement between simulations and data is worse than for the click in the previous experiment. In particular,



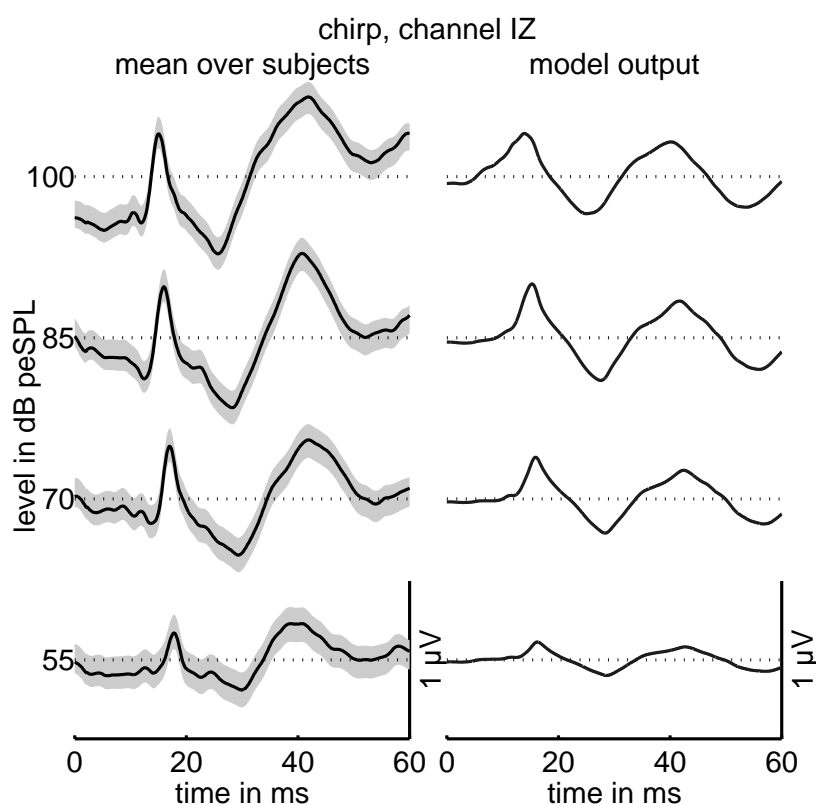
**Figure 5.4:** Amplitude (left panel) and latency data (right panel) of the potentials shown in Fig. 5.3. Closed symbols correspond to measured data and model predictions are indicated by open symbols. In the left panel the amplitude of the wave V–N19 complex is indicated by circles and amplitude of N19–P30 complex is shown with squares. In the right panel the latency of wave V, N19 and P30 is represented by circles, squares and triangles, respectively. Error bars indicate one standard error. For most latency conditions the errorbars could not be distinguished from the symbols due to the small standard errors. For better visibility, the symbols are slightly shifted along the abscissa.

the amplitudes are underestimated considerably. This is shown quantitatively in Fig. 5.6. The left panel shows the amplitudes of wave-V–N19 and N19–P30 as a function of level. The amplitudes obtained with the model (open symbols) are much too low for both the V–N19 and the N19–P30 complex. However, in contrast to the click, the saturation of the amplitude at high levels seems to be reflected in the model. The right panel of Fig. 5.6 shows the latency values for wave V, N19 and P30, as a function of level. For the waves V and N19, the predicted latency-level function is slightly flatter than the measured one. But overall, the latency values are described very well by the model.

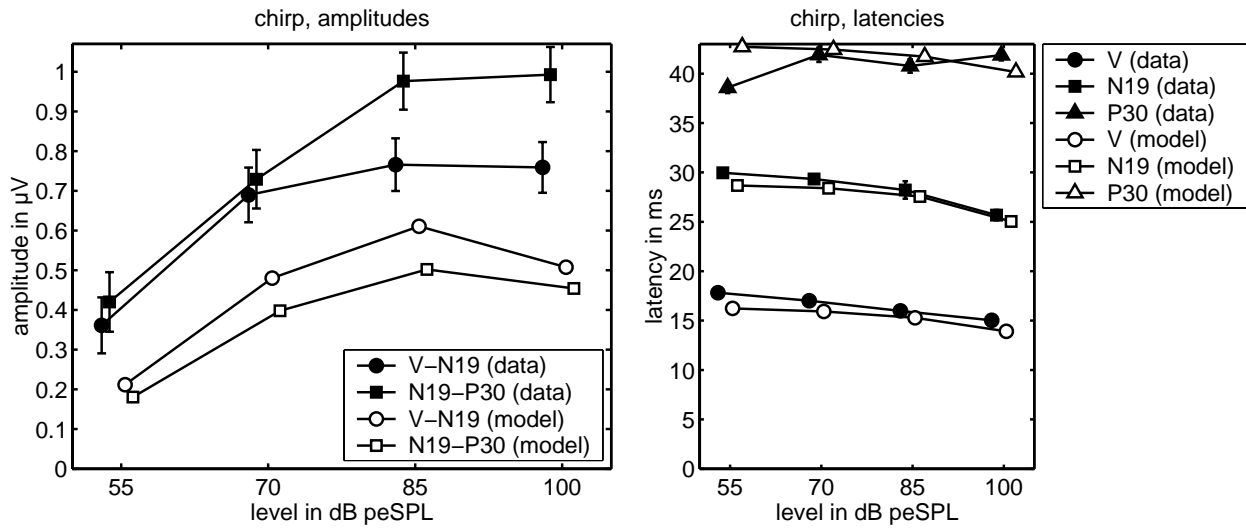
#### 5.4.2 MLRs as a function of the click rate

Grand averages of the evoked responses from all subjects are illustrated for each repetition rate in the left panel of Fig. 5.7 while the corresponding model output is shown in the right

panel. Results for different click rates are indicated on separate axes along the ordinate. The gray area in the left panel indicates  $\pm 3$  standard errors. For click rates of 16 Hz and above, the basic properties of the measured waveforms are accounted for reasonably well by the model. However, particularly for 40 Hz, but also for higher rates, the simulated response amplitudes are below the values obtained in the measured data. Fig. 5.8 (left panel) shows the amplitude of the FFT bin that corresponds to the click rate, as a function of the click rate. The amplitude of the measured response (closed circles) decreases with increasing click rate, has a minimum at 16 Hz, increases to its maximum at 40 Hz and then decreases again. The corresponding simulated response (open circles) shows a very similar pattern but with its maximum at 33 Hz (rather than 40 Hz). Furthermore, below 16 Hz, the model essentially produces a flat response amplitude. In some conditions (4, 8 and 40 Hz), the amplitude is clearly smaller than in the data. However, the main characteristics are reflected in the predictions, even if sometimes less pronounced than in the data. The right panel of Fig. 5.8 shows the corresponding phase values as a function of click rate. The model fails to predict the right phase at 4, 8 and 260 Hz, but in general, the phase behaviour of the predicted responses matches that of the recorded responses quite well.



**Figure 5.5:** Left panel: chirp evoked MLR averaged across all five subjects. Right panel: corresponding model predictions. All parameters as in Fig. 5.3

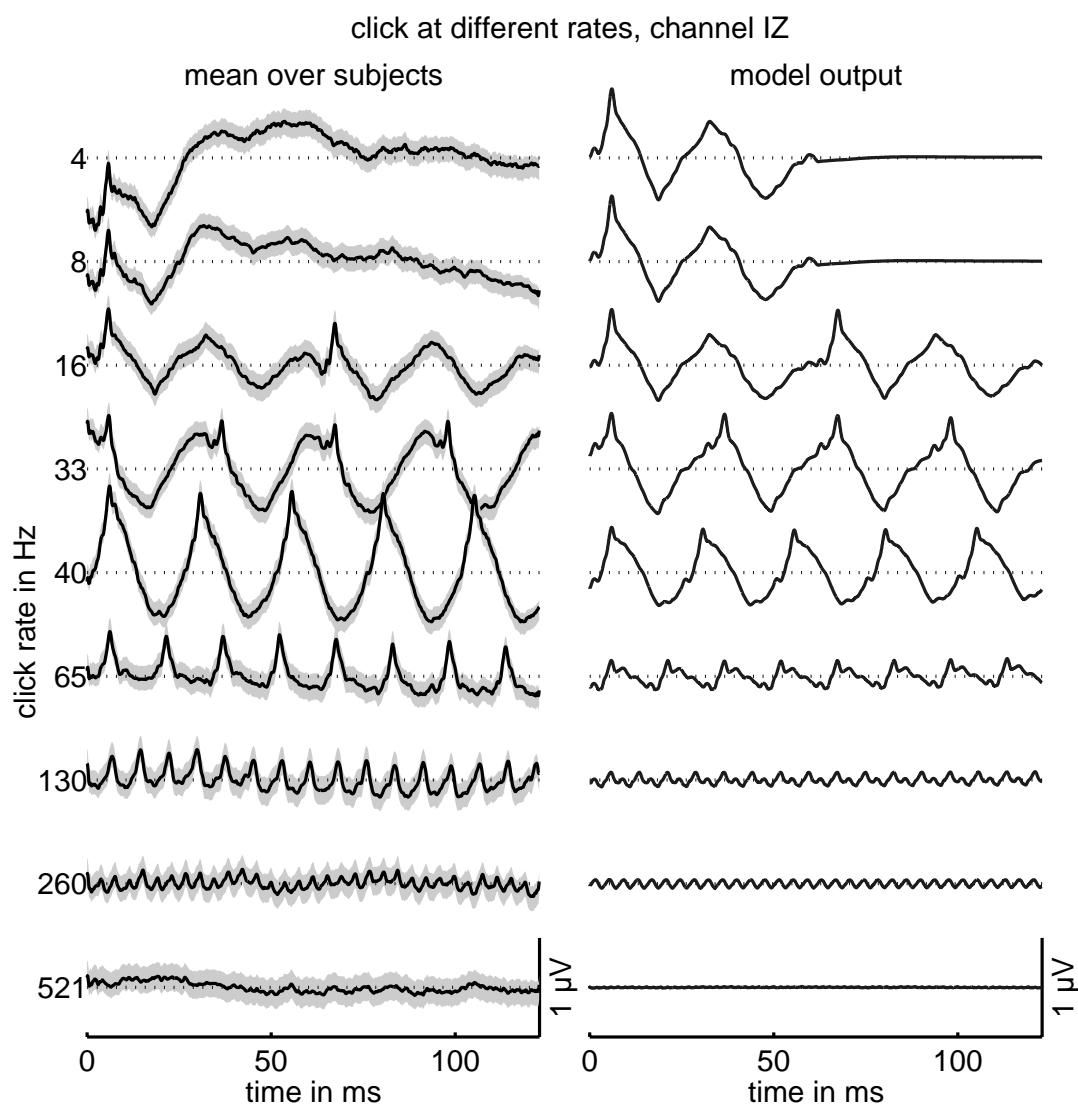


**Figure 5.6:** Amplitudes (left panel) and latencies (right panel) of the chirp-evoked potentials shown in Fig. 5.5. Closed symbols correspond to measured data, while model predictions are indicated by open symbols. Parameters as in Fig. 5.4.

These results are compatible with results from the literature: for example, Galambos *et al.* (1981) used clicks with rates in the range from 10 to 55 Hz (in steps of 5 Hz). They found the amplitude maximum always in the 35 to 45-Hz range, with the mean data showing a peak at 40 Hz. A study of Azzena *et al.* (1995) used clicks with rates of 7.9, 20, 30, 40, 50 and 60 Hz to evoke steady-state responses (SSR). They found the highest amplitude between 30 and 50 Hz. Their mean data revealed a peak at 40 Hz. As in the present study, Azzena *et al.* (1995) investigated the phase component as a function of the click rate. Their results were very similar as those described in the present study. They tried to predict their SSR data for click rates between 30 and 60 Hz by superimposing MLRs to “single” clicks at suitable time intervals and found a similar mismatch between the simulated and the measured data.

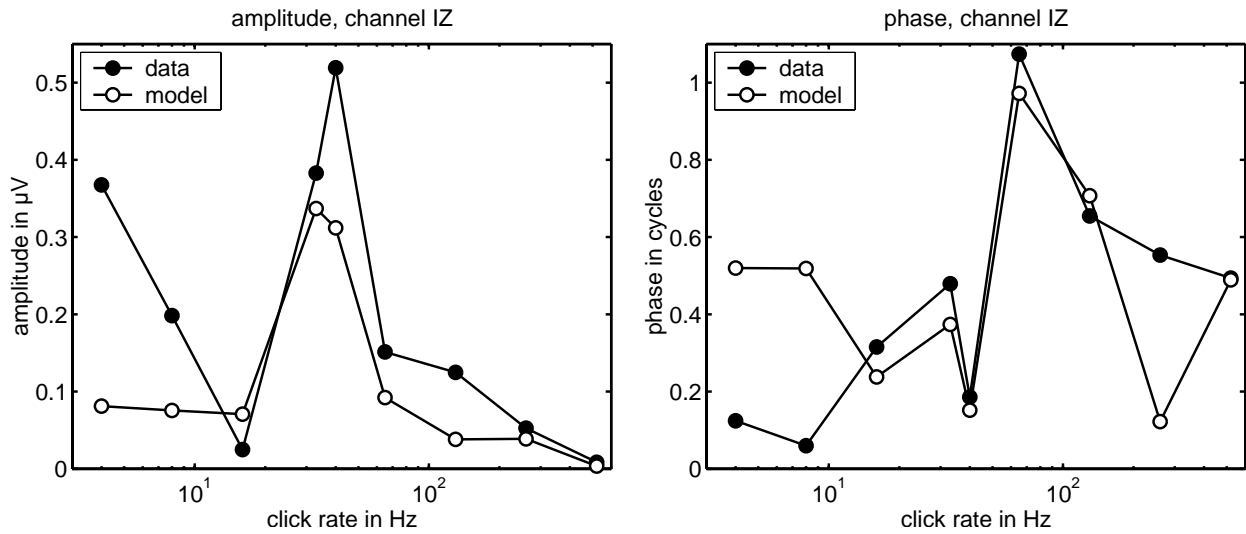
## 5.5 Discussion

In the present study, the ABR model by Dau (2003) was extended to MLR. While the experimental results obtained for click stimuli could be described reasonably well, unfortunately, this was not found for the chirp. In the data, the chirp led to larger responses than the click while this effect was clearly less pronounced in the simulations. The model systematically underestimates the amplitudes for the chirp stimulation. The good results for the click



**Figure 5.7:** Left panel: auditory evoked potentials averaged across all five subjects (channel IZ, bandpass filtered with cut-off frequencies of 2 and 1500 Hz). The potentials were elicited by a 100- $\mu$ s click at a constant level of 100 dB peSPL. Stimulus presentation rate varied between 4 and 521 Hz as indicated along the ordinate. The gray area indicates  $\pm 3$  standard errors. The right panel shows the corresponding model output.

MLRs were not very surprising, at least at the higher stimulation levels, since the UR was the result of a deconvolution of click data at a high level. But the click evoked responses are known to reflect activity from more basal, high-frequency regions of the cochlea (e.g., Neely *et al.*, 1988). In contrast, the chirp stimulus is assumed to produce synchronous discharges along the length of the human cochlear partition. The additional contributions mainly from



**Figure 5.8:** Amplitude (left panel) and phase (right panel) of the potentials shown in Fig. 5.7, plotted against the click rate. Stimulation parameters as in Fig. 5.7. Amplitude and phase of the FFT bin corresponding to the click rate were measured. Open and closed symbols correspond to model output and measured data, respectively.

low-frequency fibers lead to increased evoked potentials (Dau *et al.*, 2000; Bell *et al.*, 2002a,b; Rupp *et al.*, 2002; Wegner and Dau, 2002). Thus the chirp MLR may represent a better choice than the click stimulus to obtain the UR. In additional simulations (not shown), it was tested whether such a UR would improve the predictions and could account for both click and chirp data. However, in this case, results were better for the chirp but the simulated click response amplitudes were clearly too large. Thus, the use of the chirp-based UR leads to a similar mismatch between the predicted amplitudes of click and chirp potentials.

Dau (2003) used a chirp with a flat temporal envelope (and not a flat magnitude spectrum as in the present study) and found a good agreement between model predictions and data. Due to the flat waveform, the chirp in Dau (2003) has much more low-frequency energy than the flat-spectrum chirp used in the present study. Therefore, his results may not be directly comparable to the results presented here. However, it is possible that the ABR model would have the same difficulties if the flat-spectrum chirp was used. Thus, there might be a problem with the specific assumptions made in the preprocessing model that is responsible for the deviations – and not a problem with the principal modeling approach described here. The latter argument might be supported by the modeling results of Rupp *et al.* (2002). They also used chirp stimuli with a click-like spectrum to record middle-latency auditory evoked



fields (MAEF). They used the Auditory Image Model (AIM; [Patterson \*et al.\*, 1992, 1995](#)) to calculate the neural activity pattern (NAP). The BM stage of the model consists of a one-dimensional transmission-line filterbank. They derived different URs independently for the different stimuli they used (a click, two different chirps and their time-reversed waveforms) and compared them with each other. They found the UR to be very similar in all conditions. Thus, it appears that the AIM model predicts the delay line characteristic of the BM in a more realistic way than does the model of [Heinz \*et al.\* \(2001\)](#) which is based on (level dependent) gammatone filters.

Nonetheless, the MLR model presented here accounts for some of the main effects of the MLR as a function of the click rate. The predicted response amplitude shows a clear maximum around 32 and 40 Hz, which is very similar as in the data. However, the model underestimates the amplitude for the click at the rate of 40 Hz. This result is, in principle, comparable with the results of the [Azzena \*et al.\* \(1995\)](#) study. They investigated the mechanisms underlying the generation of the 40-Hz steady-state response (SSR) by recording click-evoked responses with different click rates. They tried to predict each response by superimposing MLRs at suitable time intervals. They concluded that (i) a model based on linear addition of transient MLRs is not able to adequately predict their results, and that (ii) other mechanisms related to the recovery cycle of the activated system play an important role in the response generation. Such recovery cycles might be assumed to be included within the MLR model presented here, because the AN model of [Heinz \*et al.\* \(2001\)](#) considers adaptation effect at least at a peripheral level. However, these are apparently not sufficient to account for the response behaviour in the measured data.

The main question of the present study was to investigate to what extent middle-latency responses can be described quantitatively using a simple linear system's "black box" approach with only very few assumptions. One assumption was that the "driving" neural excitation function, the instantaneous discharge rate functions, are approximately given by the functions known at the level of the auditory nerve. All nonlinearity within the entire modeling approach was considered to be reflected within the auditory periphery (i.e., effects of stimulation level and peripheral adaptation). Some of the key observations in the data could indeed be described well by the model, such as the general amplitude behavior at different repetition rates and, in particular, the right latency values for the click and the chirp. Many details, however, which are of particular interest such as the well known strong

response at and around the repetition rate of 40 Hz, cannot be fully explained. The origin of this strong response can therefore most likely not be explained by linear superposition of single MLR responses but need additional processes not considered in the model. Nonetheless, overall the present model might serve as a very useful tool since it can be applied to any stimulus configuration of interest. Also, the model may be applied to any form of simulated cochlear hearing loss in order to understand the effects of hearing impairment on evoked potential generation.

## 5.6 Summary and conclusions

- A model for the generation of MLR was presented which is based on the ABR model by [Dau \(2003\)](#). The model uses the concept that evoked potentials can be described by convolution of an elementary unit waveform (unitary response) with the instantaneous discharge rate function for the corresponding unit ([Goldstein and Kiang, 1958](#)). To calculate the instantaneous discharge rate functions, the AN model developed by [Heinz \*et al.\* \(2001\)](#) was used. The summed excitation across frequency was convolved with a unitary response, which was assumed to reflect contributions from different cell populations along the auditory pathway. The unitary response was derived by deconvolution of high-level click-evoked MLR.
- Predicted potential patterns were compared with corresponding experimental data for clicks and chirps at different stimulus levels, and for clicks at a wide range of different repetition rates. The main characteristics in the amplitude and latency behavior as a function of stimulation level were reflected in the simulations. Also, as a result of the linear superposition of the “single” MLRs, the model predicts an increased response amplitude for repetition rates near 40 Hz which is in qualitative agreement with the experimental data.
- However, several discrepancies between model predictions and experimental data were observed. The model generally underestimates the AEP amplitudes evoked by the flat-spectrum chirp. The reason for this discrepancy may be the use of a gammatone filterbank within the BM stage of the AN model by [Heinz \*et al.\* \(2001\)](#). The group delays of the gammatone filters may be less realistic than those obtained within the transmission-line

---

model. Also, the model does not fully account for the strong response amplitudes at and around repetition rates of 40 Hz observed in the data. Thus, the components contained in the linear system's approach of the present study are not sufficient and additional processes need to be included.

- Nevertheless, the simple generation model may serve as an estimate of MLRs and can be applied to any stimulus configuration of interest. For example, it will be interesting to also investigate “classical” steady-state responses such as envelope-following responses (EFR) or amplitude modulation following responses (AMFR) with this approach. Also, the model can be applied to any form of simulated cochlear hearing loss in order to understand the effects of hearing impairment on evoked potential generation.



# Chapter 6

## Summary and conclusions

Within the literature it is generally assumed that the conventional auditory brainstem responses (ABR) are an electrophysiological event evoked by onset or offset of an acoustic stimulus. Therefore and due to its wide spectral spread, the click is often thought of as the ideal stimulus for eliciting ABR. However, from basilar membrane mechanics it is known that the click stimulus is subject to temporal dispersion on the basilar membrane, resulting in a loss of synchronization of the auditory-nerve fiber activity along the cochlear partition. Thus, the click-evoked ABR mainly reflects contributions from more synchronous basal regions of the cochlea and not from the entire cochlea. Consequently, the click might not be the optimal stimulus for recording ABR.

Within the present thesis, the role of cochlear processing for the formation of ABR and MLR was investigated. The main emphasis was put on the development and evaluation of an optimized stimulus that is based on the concept of compensation for travel time differences along the basilar membrane. Since low-frequency tones require more time to reach their places of maximum displacement than high-frequency tones, an optimized stimulus must be a rising frequency chirp to allow compensation for the dispersion and to theoretically produce synchronous discharges of auditory-nerve fibers along the whole length of the cochlear partition. Different strategies were used in this thesis to generate such chirp stimuli and a model was presented that quantitatively tested the idea underlying the generation of these chirps.

In chapter 2, an optimized chirp based on a human cochlea model was introduced. The equations defining the chirp reflect the inverse of the delay-line characteristic of the cochlear

partition within the linear cochlea model by [de Boer \(1980\)](#). In ABR measurements with clicks, optimized chirps and temporally reversed chirps, it was shown that the chirp evokes a significantly larger wave-V amplitude than the corresponding click and the temporally reversed chirp. Additional experiments showed that this effect is due to the inclusion of activity from lower-frequency regions. Therefore, the conventional ABR should not be considered as an electrophysiological event purely evoked by onset or offset of an acoustic stimulus. Instead, an appropriate temporal organization of the stimulus, determined by basilar membrane traveling wave properties, is important to ensure that all frequency regions contribute to the evoked potential. The chirp may therefore be of clinical use in assessing the integrity of the entire human cochlea.

The usefulness of the chirp for retrieving frequency-specific information was investigated in [chapter 3](#). Frequency-specific evoked ABR are an important clinical tool for estimating frequency-specific hearing especially at low frequencies. Two different strategies were used to compare the chirp stimulus with corresponding standard stimuli: the first one compared chirp-evoked ABR with click-evoked ABR, both in the presence of different high-pass and notched-noise maskers. The second strategy used a low-frequency chirp and a tone pulse (with a comparable duration and magnitude spectrum) to evoke ABR at different stimulation levels. It was observed that the increased synchrony obtained with the chirp relative to the click stretches over the entire frequency region. Thus, the chirp may be particularly interesting for clinical use in the low-frequency region. It was also shown that the narrow-band chirp evokes a larger response than the corresponding tone pulse. An auditory nerve (AN) model ([Heinz \*et al.\*, 2001](#)) was used to qualitatively explain the differences between the click and the (broadband) chirp as well as between the tone-pulse and the low-frequency chirp. The results further demonstrate the importance of cochlear processing for the formation of ABR.

In [chapter 4](#), two alternative chirp stimuli were developed. In contrast to the model-based chirp from [chapter 2](#), the equations defining these two chirps are based on experimental data. In one case, the basilar membrane group delay was estimated from stimulus frequency otoacoustic emissions (SFOAE) in human subjects ([Shera and Guinan, 2000](#)). In the other case, the basilar membrane group delay estimates were derived from tone-pulse-evoked ABR wave-V latency data in humans ([Gorga \*et al.\*, 1988](#); [Neely \*et al.\*, 1988](#)). Using these estimates, derived for a wide range of frequencies and levels, a level-dependent ABR-based chirp

---

stimulus was developed. ABR evoked by all three chirps were compared with click-evoked responses. It was shown that all chirps produced a larger response amplitude than the click. Differences between the chirp responses vary with the stimulation level. Overall, the level-dependent chirp was found to produce the largest response amplitudes. Its advantage is particularly large at low stimulation levels which makes this chirp particularly interesting for clinical applications such as the objective assessment of hearing thresholds.

In chapter 5, a model for the generation of middle-latency responses (MLR) was introduced, which itself is an extension of a model for ABR generation (Dau, 2003). It uses the concept that evoked potentials can be described as a convolution of the instantaneous discharge rate function of a unit with its elementary unit waveform (unitary response). The instantaneous discharge rate functions were calculated with a computational AN model. The unitary response was derived by deconvolving high-level click-evoked MLR data with the simulated summed AN activity for the click. This unitary response was used for any input stimulus at any level, implying linearity at this stage of processing. All nonlinearity in the model is assumed to be restricted to the processing of the stimulus-dependent rate functions in the AN. Predicted potential patterns using clicks and chirp stimuli at different levels were compared with corresponding experimental data. Additional experiments were performed, using clicks at different repetition rates, in order to investigate steady state responses (SSR). The main characteristics of the data were reflected in the model predictions. However, several discrepancies between the data and the predictions were observed, which may, at least partly, be caused by the use of the gammatone filterbank within the AN model. Corresponding simulations on the basis of a transmission line filterbank may lead to better results. Also, future studies are needed that investigate more closely the mechanisms that are responsible for the large amplitudes derived at repetition rates near 40 Hz. Overall, the model could be an interesting tool for investigation, since it can be applied to any stimulus configuration of interest. The model might also be applied to simulate cochlear hearing losses to get a better understanding of the effects of hearing impairment on evoked potential generation.

To summarize, in the present thesis the importance of considering the (nonlinear) effects of basilar-membrane traveling wave and AN-processing for the formation of ABR and MLR was demonstrated. Using optimized chirp stimuli, it was shown that the AEP do not necessarily represent a response to the onset or offset of acoustic stimuli. Instead, it was shown

that it is important to achieve a high neural synchronization at the level *after* cochlear preprocessing in order to generate a large far-field response. One main consequence from this study is that the chirp stimulus may have large impact in clinical diagnostics of the human auditory system, since it causes an increase of neural synchrony over the entire frequency region in comparison to standard stimuli (e.g., the click). This leads to an improved signal-to-noise ratio of the evoked potential and a better assessment of the low-frequency region. It can be expected that the use of optimized chirps will also be very valuable as objective indicator for hearing thresholds in the impaired auditory system. Another important consequence from the present study is that the hypotheses underlying the generation of the chirp can be used to explicitly test current models of human cochlear (and retro-cochlear) processing. This is of particular relevance since a “realistic” model would allow for a better interpretation of empirically obtained potential patterns whereby any stimulus configuration could be tested for any hearing loss assumed within the framework of the model. Also, if successful in predicting the potential patterns, such a model of peripheral signal processing will also be very valuable in other important applications, for example as a preprocessing stage in models of auditory perception and for automatic speech recognition.



# Appendix A

## Experiments on the correlation between psychophysical loudness and auditory brainstem responses<sup>1</sup>

### A.1 Introduction

Psychophysical loudness does not only depend on stimulus level, but also on other stimulus properties like spectral composition. For example, given a constant physical stimulus level, the perceived loudness increases with increasing bandwidth of the stimulus. This effect is called loudness summation. Present loudness models use the concept of critical bandwidth with subsequent compression to explain this effect. These processes are generally assumed to be peripherally located in the auditory system, namely within the cochlea. If this is true, it should be possible to measure the effect of loudness summation with auditory brainstem responses (ABR). However, to be able to examine these spectral effects with ABR, it is important to ensure simultaneous excitation of the different critical bands on the level of the basilar membrane. Thus an “optimal” stimulus is needed. Such a stimulus was introduced by (Dau *et al.*, 2000). Supposed there is a direct correlation between wave-V amplitude and loudness, wave-V amplitude should be constant when stimulating with “optimal” stimuli with different bandwidths and equal loudness.

---

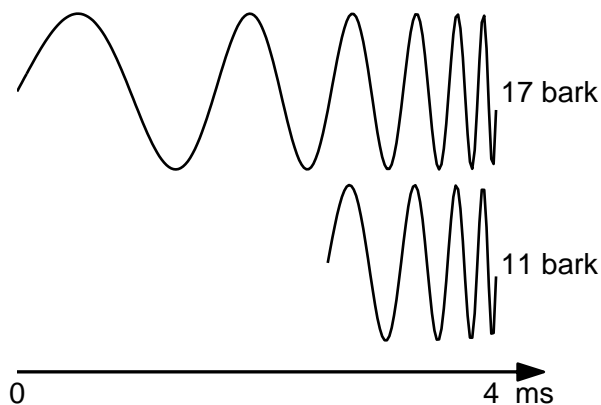
<sup>1</sup> Part of this research was presented together with Torsten Dau, Jesko Verhey and Birger Kollmeier at the “24. Deutsche Jahrestagung für Akustik” of the “Deutsche Gesellschaft für Akustik”.

The present study tests this hypothesis. Optimized chirp stimuli (Dau *et al.*, 2000) with different bandwidths were used in a psychoacoustical loudness-matching experiment. The levels of constant loudness were used to evoke ABRs. Predictions of different loudness models were compared with the results of this study.

## A.2 Methods

### A.2.1 Psychophysics

Psychophysical loudness-matching experiments were performed using a 2-AFC 1-up 1-down procedure. Experiments were performed with an “interleaved” and with a “non-interleaved” presentation of the different signal bandwidths to examine effects of stimulation paradigm (Verhey and Kollmeier, 1998). Optimized chirps, which compensate for the travel time differences on the basilar membrane (Dau *et al.*, 2000), were used as stimuli. The upper cut-off frequency of the chirps was set to 8.5 kHz, while the lower cut-off frequencies were 0.7, 1, 1.37, 1.85, 2.5, 3.4 and 4.8 kHz. These cut-off frequencies correspond to bandwidths of 3, 5, 7, 9, 11, 13, 15 and 17 bark. Due to the properties of the optimized chirp, the different bandwidths result in different stimulus durations (Fig. A.1). The chirps were presented as trains of 10 chirps with a duration of 500 ms and as single chirps. The “chirp rate” within the trains was 20 Hz corresponding to the repetition rate in the ABR experiments. The reference stimulus in the non-interleaved loudness matching experiment was a chirp train with a bandwidth of 17 bark at a level of 47 dB SPL. For the interleaved experiment this was a chirp train with a bandwidth of 9 bark and a level of 55 dB SPL.



**Figure A.1:** Temporal course of the chirps with a bandwidth of 11 and 17 bark. Note the different durations of the stimuli, which result directly from the different bandwidths.

## A.2.2 Evoked potentials

ABR experiments were carried out with a personal computer which controlled stimulus presentation and recording of evoked potentials. A DSP-card (Ariel DSP32C) converted the digitally generated stimuli (25 kHz, 16 bit) to an analog waveform. The output of the DSP card was connected to a digitally controlled audiometric amplifier which presented the stimulus through an insert earphone (Etymotic Research ER-2) to the ear of the subject.

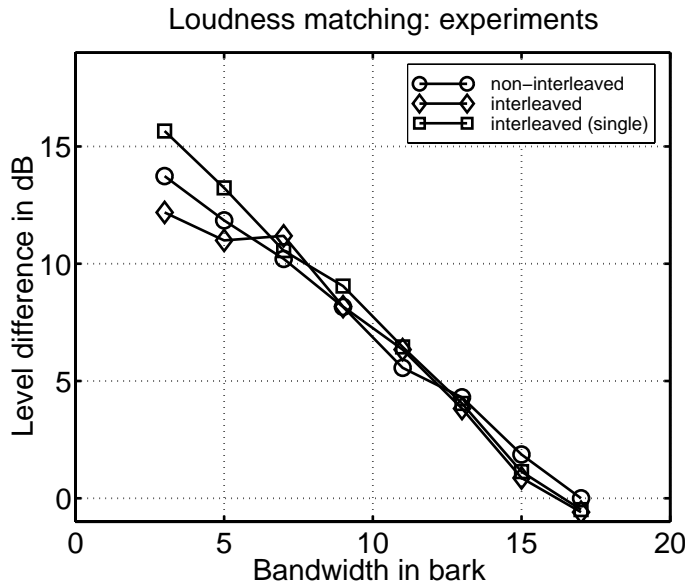
Electroencephalic activity was recorded from the scalp via silver/silver-chlorid electrodes, attached to the vertex (positive) and the ipsilateral mastoid (negative). The forehead served as the site for the ground electrode. Interelectrode impedances were maintained below 5 k $\Omega$ . Responses were amplified (80 dB) and bandpass filtered (95–1640 Hz, 6 dB/oct) with a commercially available ABR preamplifier (Hortmann Neurootometrie). Extra amplification (Kemo VBF/40) was used to match the optimum range for the A/D converter of the DSP card. This amplification was in the range from 10 to 16 dB, resulting in a total amplification of 90–96 dB. The amplified signal was digitized by the DSP card (25 kHz, 16 bit), which also performed artifact rejection and signal averaging. Responses were recorded for 16 ms following stimulus offset.

ABR elicited by the bandlimited chirps were recorded at equal loudness levels. In addition, corresponding responses were recorded at a constant stimulus amplitude of 54 dB peSPL for each bandwidth condition.

## A.3 Results

### A.3.1 Psychoacoustical experiments

Figure A.2 shows the results from the psychoacoustical loudness matching experiments. It shows the level difference with respect to the reference stimulus which results in equal loudness of the stimuli as a function of the stimulus bandwidth. For better comparison with the non-interleaved data, the interleaved data are shifted by 8 dB to higher values (according to the level difference of the reference stimuli). The functions are the mean of all four subjects, who participated in this study. The standard errors within and between subjects were both below 3 dB.



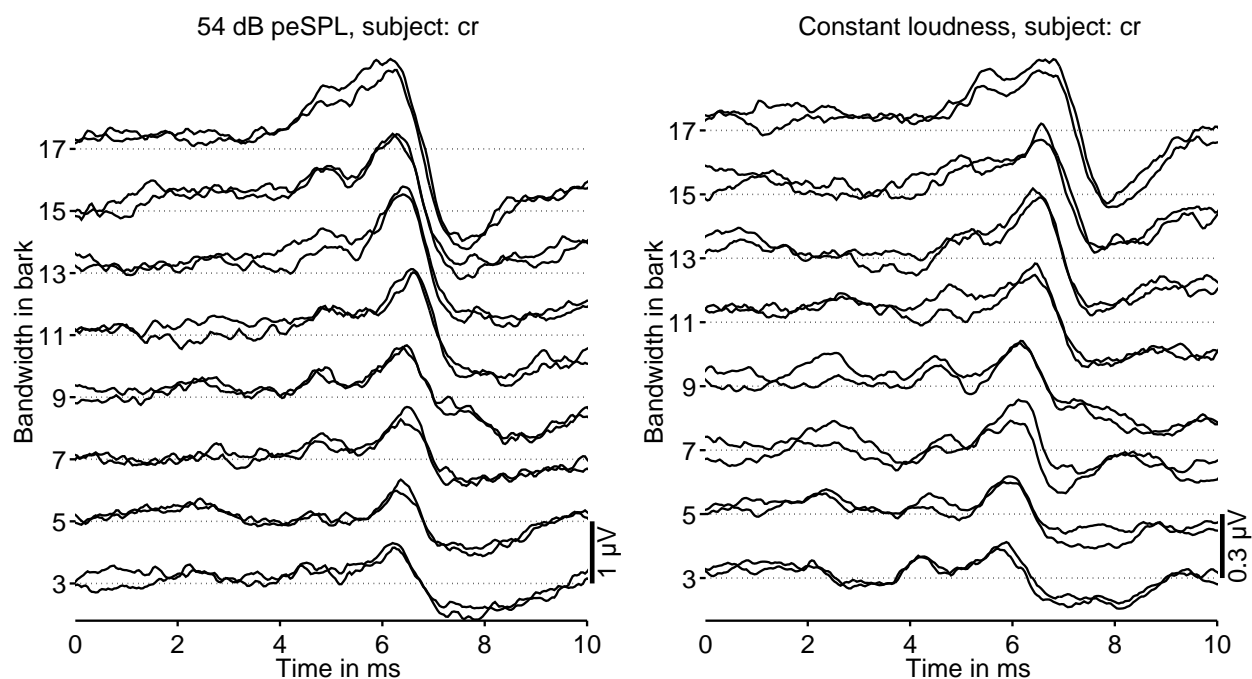
**Figure A.2:** Results from the psychoacoustical loudness matching experiments. The abscissa shows the bandwidth of the chirp stimuli and the ordinate shows the level difference needed to perceive the stimulus as loud as the reference stimulus. Results for the interleaved experiments were shifted by about 8 dB to higher values for better comparison with the non-interleaved data. The plotted functions represent the mean over four subjects.

Figure A.2 shows that – independent of the measurement procedure – with decreasing stimulus bandwidth the stimulus level must be increased to perceive the stimuli as equally loud. For example, the stimulus with the smallest bandwidth (3 bark) must be presented with an about 14 dB higher level (mean over all measurements) to be perceived as loud as the stimulus with the largest bandwidth (17 bark). This effect is comparable with results in the literature (e.g., Zwicker and Fastl, 1990). The different stimulation paradigms result in comparable level differences. Only small differences (up to about 3 dB) at small bandwidths (3 and 5 bark) could be observed between the different stimulation paradigms.

### A.3.2 Evoked potentials

Results of the ABR measurements are shown in Fig. A.3 for one exemplary subject. As expected, wave-V amplitude increases with constant stimulus amplitude and increasing bandwidth (Fig. A.3, left panel). But stimulation with chirps of constant loudness (and increasing bandwidth) results in increasing wave-V amplitudes, too (left panel of Fig. A.3). However, this increase is much smaller than the one for the constant stimulus amplitude.

Figure A.4 shows the mean results for the ABR measurements. The left panel shows the mean wave-V amplitude for stimulation with constant stimulus amplitude as a function of stimulus bandwidth. Due to the large variation of wave-V amplitudes across subjects, the evoked potentials were normalized before averaging: the largest individual wave-V amplitude

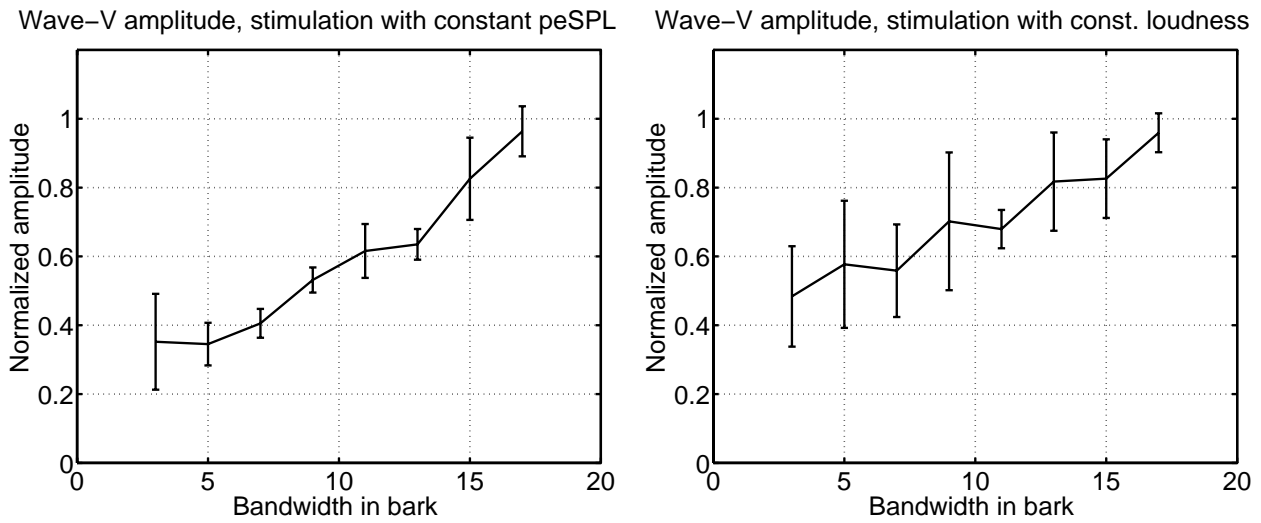


**Figure A.3:** ABR results for subject *cr*. Left column: ABR evoked by chirps with varying bandwidth and equal amplitude (54 dB peSPL). Right panel: ABR evoked by chirps with varying bandwidth and equal loudness. Notice the different scaling of the ordinates in the left and right panel.

was set to one. As in Fig. A.3, wave-V amplitude increases with increasing bandwidth. This indicates some kind of integration across frequencies. However, this increase could also be seen for the stimulation with constant loudness (right panel of Fig. A.4), but in this case the increase is a bit smaller: in the first case the normalized amplitude increases from about 0.35 to 0.96 while in the latter case it increases from about 0.5 to 0.96. None of the subjects showed a constant wave-V amplitude for stimulation with constant loudness.

## A.4 Discussion

The results of the ABR experiments show that loudness compensation results in smaller wave-V amplitude differences (compared to stimulation with constant stimulus amplitude), but amplitudes were *not* equal. Thus stimuli with smaller bandwidths must be presented with even higher levels to get equal wave-V amplitudes.

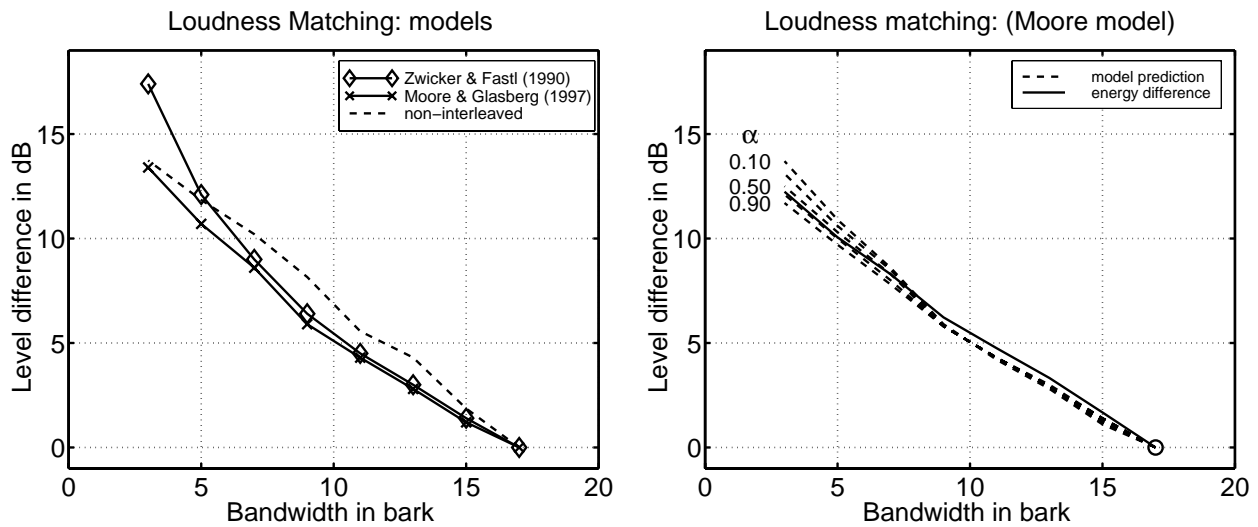


**Figure A.4:** Normalized wave-V amplitude as a function of chirp bandwidth (mean over four subjects). The left panel shows the results for the stimulation with constant stimulus amplitude. Results for the stimulation with constant loudness are shown in the right panel.

Psychoacoustic models use the concept of critical bandwidth to explain loudness summation: the excitation within each critical band is determined and then compressed (by raising it to the power of  $\alpha < 1$ ). The compressed excitations are summed up across all critical bands. Assuming that the wave-V amplitude can be explained by a similar concept, there must be some stronger compression at the level of the brainstem to be able to explain the data presented here.

To test, whether the bandwidth dependency of the wave-V amplitude can be explained with a loudness model using a higher compression, the model must meet some requirements. First of all the model should be able to predict the psychophysical experiments. Another requirement is that the model prediction of the experiment must critically depend on the compression parameter (the exponent  $\alpha$ ).

The left panel of Fig. A.5 shows two different model predictions (solid lines) of the loudness matching experiment (dashed line). Both models (Moore and Glasberg, 1997; Zwicker and Fastl, 1990) were developed for loudness predictions of stationary signals. They are able to predict the results of the loudness matching experiment (for transient signals) quite well. However, while the model from Moore and Glasberg (1997) predicts the maximal effect correctly, the model from Zwicker and Fastl (1990) fails to do this. Therefore the first model is used in the right panel of Fig. A.5 to test how critically the prediction depends on



**Figure A.5:** Left panel: simulation of the psychoacoustical loudness matching experiment with different loudness models (Zwicker and Fastl, 1990; Moore and Glasberg, 1997). The right panel illustrates the variation of the prediction of the model from Moore and Glasberg (1997) with the compression parameter  $\alpha$  (dashed lines). The solid line represents the level differences which result in equal energy of all stimuli.

the compression parameter  $\alpha$ . Despite a large variation of  $\alpha$  (from 0.1 to 0.9), the model prediction maximally varies by about 2 dB. However, a simple compensation for the energy difference between the different stimuli (solid line in the right panel of Fig. A.5) explains the main trend of the data. These energy differences were caused by the different durations of the chirps. Thus an effect of loudness summation – which is not just energy summation – was hardly found within the experiments presented here.

The concept which loudness models use to explain effects of loudness summation (excitation in each critical band – compression – summation) can therefore *not* be used to describe the wave-V amplitude in the experiments presented here. The results rather point to a nonlinear relationship between loudness perception and wave-V amplitude. To explain the results presented here one could think to introduce a bandwidth dependent compression, such that the compression increases with increasing bandwidth. Alternatively a time dependent compression could be used.

## A.5 Summary

- ABR evoked by optimized chirps of different bandwidths and constant stimulus amplitude (peSPL) result in increasing wave-V amplitude with increasing bandwidth.
- ABR evoked by the same chirps presented with equal loudness results in smaller increase of wave-V amplitude with increasing bandwidth.
- It is not possible to explain the relation between wave-V amplitude and chirp bandwidth with the concept of critical bands as used in loudness models.
- There is rather a nonlinear relationship between loudness and wave-V amplitude.
- To get a deeper insight, a modified concept is needed, e.g., a compression dependent on bandwidth or time.



# Appendix B

## On the relationship between auditory evoked potentials and psychophysical loudness<sup>1</sup>

### B.1 Introduction

*Wegner et al. (1998)* developed optimized chirp stimuli to examine the electro-physiological correlate of loudness summation at the brainstem level. They found that loudness does not directly relate to the auditory brainstem response (ABR) amplitude. *Nousak and Stapells (1998)* examined loudness-growth functions for tone pulses. They compared psychophysical data with amplitude data of ABR wave V and middle latency responses ( $N_a$ - $P_a$ ). Their results were similar to those obtained by *Wegner et al. (1998)* suggesting a higher compression as derived from psychophysical data and as reflected in common loudness models. Several authors stated that the cortical auditory evoked potentials (CAEP) – namely the  $N_1$ - $P_2$ -component – reflect an integration process of neural activity over several milliseconds (e.g., *Davis, 1974; Eddins, 1998*).

The present study examines whether effects of loudness summation observed in psychoacoustical experiments are reflected in CAEP. In particular, the role of synchronisation and effects of temporal integration at this stage of processing are investigated.

---

<sup>1</sup> This chapter was published together with Torsten Dau and Birger Kollmeier, see *Wegner et al. (1999)*.

## B.2 Methods

The experiments were carried out with a PC based ERA-system (ZLE-Systeme). The digitally generated stimuli (40 kHz, 12 bit) were presented through an insert ear-phone (Etymotic Research ER-2) to the subjects. Electroencephalic activity was recorded from the scalp via silver/silver chloride electrodes, attached to the vertex (positive) and the ipsilateral mastoid (negative). The forehead served as site for the ground electrode. Interelectrode impedances were maintained below 5 k $\Omega$ . Responses were amplified (88 dB) and filtered (analog: 1–100 Hz, 6 dB/oct; digitally: 16 Hz low-pass, 2nd order).

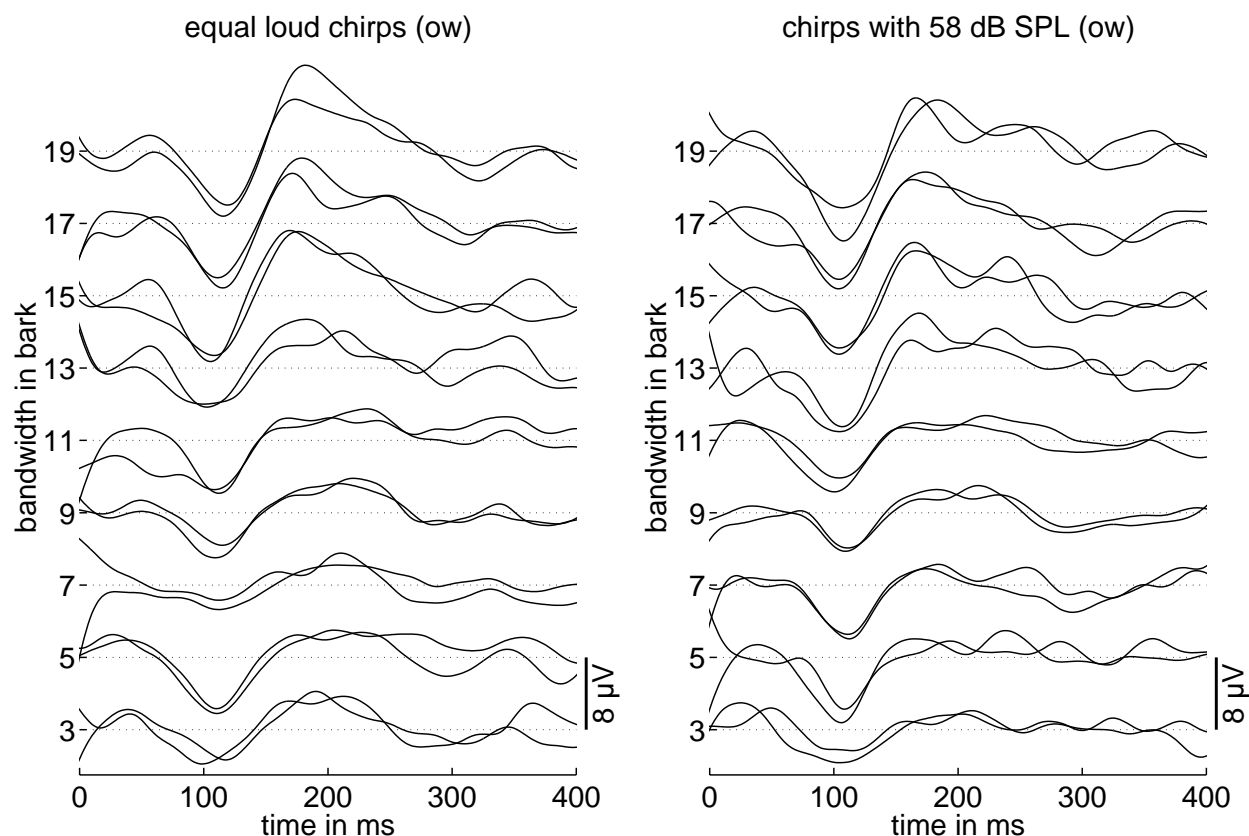
## B.3 Results

### B.3.1 Experiment 1

Psychophysical loudness-matching experiments were performed using a 2-AFC 1-up 1-down procedure with an “interleaved” presentation of the different signal bandwidths. The signals were chirps of 100 ms duration with linearly changing instantaneous frequency (Neumann *et al.*, 1994); their bandwidths varied in the range from 3 to 19 bark in steps of 2 bark. The reference stimulus was a chirp with a bandwidth of 9 bark at a level of 58 dB SPL. On average the 3-bark wide chirp was adjusted at a 23 dB higher level than the 19-bark wide chirp to be equal loud.

CAEP elicited by these bandlimited chirps were recorded. They are shown in the left panel of Fig. B.1. In addition, corresponding responses were recorded at a constant level of 58 dB SPL at each bandwidth condition, shown in the right panel of Fig. B.1.

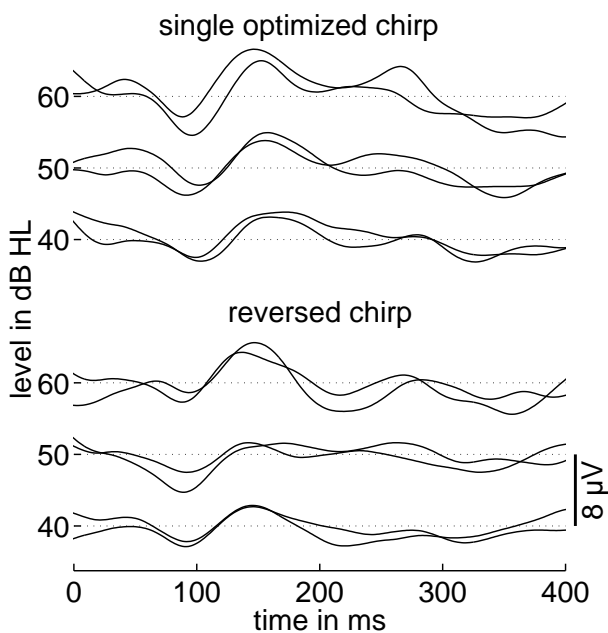
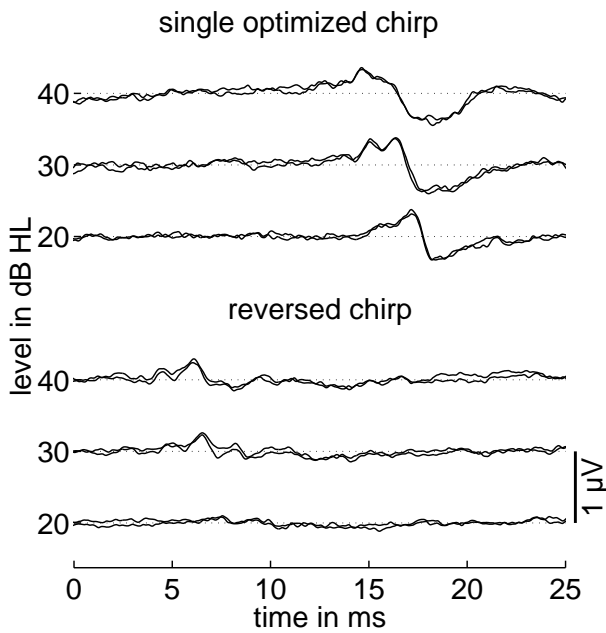
The  $N_1$ – $P_2$  amplitude of the CAEP increases with increasing chirp bandwidth (left panel). This suggests that an integration of neural activity across frequencies occurs at this level of processing. The results also show that equal loudness does not seem to be reflected in an equally large  $N_1$ – $P_2$  amplitude of the CAEP (right panel). However, only small differences in potential amplitude occur for the different conditions tested.



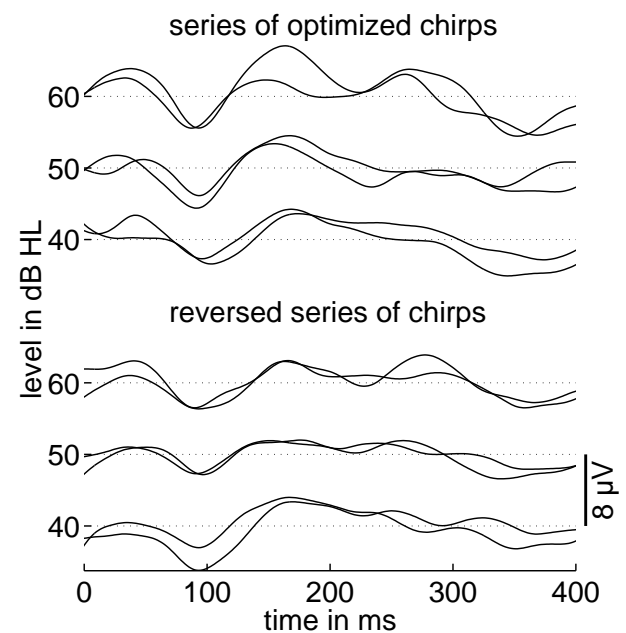
**Figure B.1:** CAEP evoked by bandlimited chirps with constant loudness (left panel) and with constant sound pressure level (right panel), respectively. The chirp duration was 100 ms. The recording time started at stimulus onset. The waveforms are the average of 32 responses for each buffer. The stimulus onset asynchrony (SOA) was 5.504 s. Subject: ow.

### B.3.2 Experiment 2

This experiment examines the role of synchronization at cortical level. CAEP evoked by single rising and falling chirps were recorded. In this case, nonlinear rising chirp stimuli were used which were developed to compensate basilar-membrane dispersion. These chirps have been shown to maximize the synchronization at brainstem level (ABR) (Dau *et al.*, 2000) and were therefore called “optimized chirps”. Fig. B.2 shows ABR (left panel) and CAEP (middle panel) obtained with single rising chirps (top) and single temporally reversed (falling) chirps (bottom panel). In addition, CAEP were recorded with a 200 ms long series of four chirps (right panel of Fig. B.2). The repetition rate was 20 Hz.



**Figure B.2:** Potentials evoked by the optimized broadband chimp stimulus and the time-reversed chimp stimulus. Upper left: ABR for stimulation levels of 20–40 dB HL. The stimulus onset asynchrony (SOA) was 50 ms. Lower left: CAEP for stimulation levels of 40–60 dB HL. Waveforms are averaged over 32 responses and the SOA was 5504 ms. Lower right: CAEP evoked by a 200-ms series of chirps at a repetition rate of 20 Hz. Stimulus level, averages and SOA as in the middle panel. Subjects: cr/os.



The data show that for the generation of CAEP neural synchronization is much less important than for the generation of ABR: in most conditions, the chimp stimuli evoke the same potential amplitude as the temporally reversed chimp stimuli. The results also show that neural activity is not temporally integrated when a series of multiple short-duration chirps is presented. However, due to the small number of averages the signal-to-noise ratio is relative small (indicated by the two different waveforms for each condition). Therefore further experiments with an even higher number of averages are needed.

### B.3.3 Experiment 3

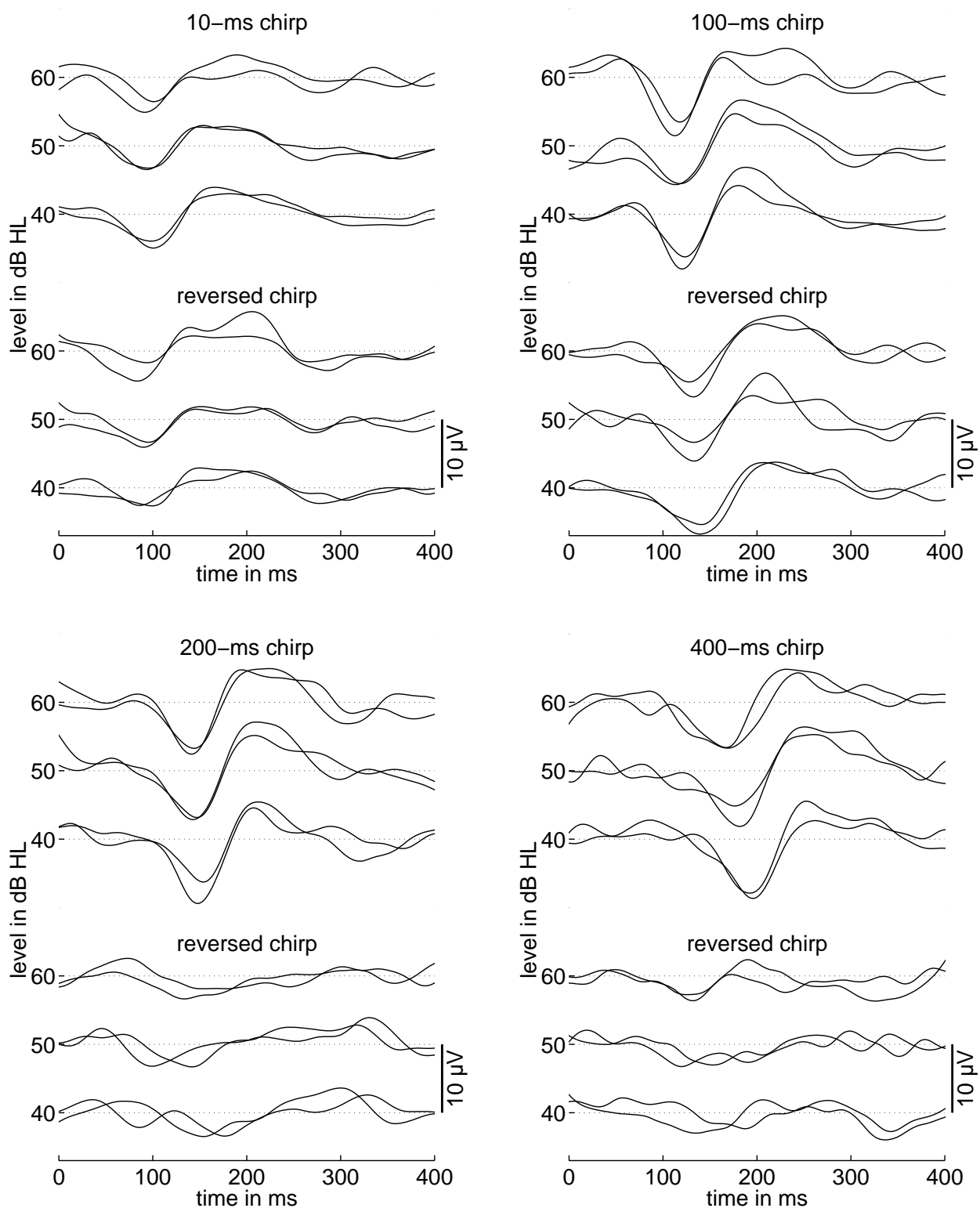
As mentioned in the introduction, there seems to be some evidence that the  $N_1$ - $P_2$ -component of the CAEP reflects an integration process of neural activity over several milliseconds (e.g., [Davis, 1974](#); [Eddins, 1998](#)). The aim of this experiment was to derive the time constant of this integration process.

The stimuli used here were chirps with linearly changing instantaneous frequency as in Exp. 1. The chirp bandwidth was 9 kHz. Four different durations were used: 10, 100, 200 and 400 ms. Each stimulus was presented with rising and with falling instantaneous frequency at levels of 40, 50 and 60 dB HL, respectively.

The data are presented in Fig. B.3. The amplitude of the  $N_1$ - $P_2$  complex increases up to a chirp length of 100 ms. No further increase can be seen for longer durations. Thus neural activity appears to be integrated for stimulus durations up to about 100 ms, whereas for longer durations, no further integration seems to occur.

## B.4 Summary and conclusion

- An integration of neural activity across frequencies occurs at the level of processing where  $N_1$ - $P_2$  amplitude of the CAEP is generated. However, loudness summation does not seem to be directly reflected in the potential amplitude.
- The fact that only small differences in potential amplitude occur for the different conditions tested in Exp. 1 may be due to the relatively high stimulation level at which neural activity possibly is near saturation. Similar experiments should be performed at lower stimulation levels.
- For the generation of CAEP, neural synchronization is much less important than for the generation of ABR.
- Neural activity at cortical level as reflected in CAEP amplitude, appears to be integrated up to about 100 ms, whereas it remains constant for longer durations.
- Further experiments are currently investigated to clarify the role of onsets and offsets for the generation of CAEP and the influence of the spectral shape of the stimulus.



**Figure B.3:** Results of Exp. 3. CAEP evoked by rising and falling linear chirps with levels of 40–60 dB HL. Four different durations were tested: 10 ms, 100 ms, 200 ms, and 400 ms. Other parameters were the same as in Fig. B.2 (right panel).

# References

Abdala C. and Folsom R.C. (1995a):

“Frequency contribution to the click-evoked auditory brain-stem response in human adults and infants,” *J. Acoust. Soc. Am.* **97**(4), 2394–2404.

Abdala C. and Folsom R.C. (1995b):

“The development of frequency resolution in humans as revealed by auditory brain-stem response recorded with notched-noise masking,” *J. Acoust. Soc. Am.* **98**(2 Pt. 1), 921–930.

Azzena G.B., Conti G., Santarelli R., Ottaviani F., Paludetti G., and Maurizi M. (1995):

“Generation of human auditory steady-state responses (SSRs). I: Stimulus rate effects,” *Hear. Res.* **83**(1–2), 1–8.

Batra R., Kuwada S., and Maher V.L. (1986):

“The frequency-following response to continuous tones in humans,” *Hear. Res.* **21**(2), 167–177.

Beattie R.C. and Kennedy K.M. (1992):

“Auditory brainstem response to tone bursts in quiet, notch-noise, high-pass noise, and broadband noise,” *J. Am. Acad. Audiol.* **3**(5), 349–360.

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.

Beattie R.C., Franzone D.L., and Thielen K.M. (1992):

“Effects of notch noise bandwidth on the auditory brainstem response to clicks,” *J. Am. Acad. Audiol.* **3**(4), 269–274.

Beattie R.C., Garcia E., and Johnson A. (1996):

“Frequency-specific auditory brainstem responses in adults with sensorineural hearing loss,” *Audiology* **35**(4), 194–203.

von Békésy G. (1960):

*Experiments in Hearing* (McGraw-Hill, New York).

Bell S.L., Allen R., and Lutman M.E. (2002a):

“Optimizing the acquisition time of the middle latency response using maximum length sequences and chirps,” *J. Acoust. Soc. Am.* **112**(5), 2065–2073.

Bell S.L., Allen R., and Lutman M.E. (2002b):

“An investigation of the use of band-limited chirp stimuli to obtain the auditory brainstem response,” *Int. J. Audiol.* **41**(5), 271–278.

de Boer E. (1975):

“Synthetic whole-nerve action potentials for the cat,” *J. Acoust. Soc. Am.* **58**(5), 1030–1045.

de Boer E. (1980):

“Auditory physics. Physical principles in hearing theory I,” *Phys. Rep.* **62**, 87–174.

Brinkmann R. and Scherg M. (1979):

“Human on- and off-potentials of the brainstem. Influence of stimulus envelope characteristics,” *Scand. Audiol.* **8**(1), 27–32.

Brugge J.F., Anderson D.J., Hind J.E., and Rose J.E. (1969):

“Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds,” *J. Neurophysiol.* **32**(3), 386–401.

Bunke D., von Specht H., Mühler R., Pethe J., and Kevanishvili Z. (1998):

“Der Einfluß der Reizanstiegszeit und der Hochpaßmaskierung auf die frühen auditorisch evozierten Potentiale,” *Laryngorhinootologie* **77**(4), 185–190.

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.



van Campen L.E., Hall J.W., and Grantham D.W. (1997):

“Human offset auditory brainstem response: Effects of stimulus acoustic ringing and rise-fall time,” *Hear. Res.* **103**(1–2), 35–46.

Carney L.H. (1993):

“A model for the responses of low-frequency auditory-nerve fibers in cat,” *J. Acoust. Soc. Am.* **93**(1), 401–417.

Coats A.C., Martin J.L., and Kidder H.R. (1979):

“Normal short-latency electrophysiological filtered click responses recorded from vertex and external auditory meatus,” *J. Acoust. Soc. Am.* **65**(3), 747–758.

Conijn E.A.J.G., 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.

Conijn E.A.J.G., Brocaar M.P., and van Zanten G.A. (1992a):

“Low-frequency specificity of the auditory brainstem response threshold elicited by clicks masked with 1590-Hz high-pass noise in subjects with sloping cochlear hearing losses,” *Audiology* **31**(5), 272–283.

Conijn E.A.J.G., Brocaar M.P., van Zanten G.A., and van der Drift J.F.C. (1992b):

“Comparison between the frequency specificities of auditory brainstem response thresholds to clicks with and without high-pass masking noise,” *Audiology* **31**(5), 284–292.

Conijn E.A.J.G., Brocaar M.P., and van Zanten G.A. (1993):

“Frequency-specific aspects of the auditory brainstem response threshold elicited by 1000-Hz filtered clicks in subjects with sloping cochlear hearing losses,” *Audiology* **32**(1), 1–11.

Cooper N.P. and Rhode W.S. (1997):

“Mechanical responses to two-tone distortion products and in apical and basal turns of the mammalian cochlea,” *J. Neurophysiol.* **78**, 261–270.

Dau T. (2001):

“The importance of basilar-membrane and auditory-nerve processing for the formation of auditory brainstem responses,” in *24th meeting of the Association for Research in Otolaryngology* p. 195.

Dau T. (2003):

“The importance of cochlear processing for the formation of auditory brainstem and frequency following responses,” *J. Acoust. Soc. Am.* (in press).

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. (1974):

“Relations of peripheral action potentials and cortical evoked potentials to the magnitude of sensation,” in *Sensation and measurement*, edited by H.R. Moskowitz, B. Scharf, and J.C. Stevens (Reidel Publishing, Boston), pp. 37–47.

Davis H. (1976):

“Principles of electric response audiometry,” *Ann. Otol. Rhinol. Laryngol*, Suppl. 3 **28**(3), 1–96.

Davis H. and Hirsh S.K. (1976):

“The audiometric utility of brainstem responses to low-frequency sounds,” *Audiology* **15**(3), 181–195.

Debruyne F. (1982):

“Frequency specificity and on-effect in brainstem electric response audiometry,” *J. Otolaryngol.* **11**(4), 267–270.

Debruyne F. and Forrez G. (1982):

“On-effect in brainstem electric response audiometry. Consequences for the use of tonebursts,” *J. Otorhinolaryngol. Relat. Spec.* **44**(1), 36–42.

Dolan D.F., Teas D.C., and Walton J.P. (1983):

“Relation between discharges in auditory nerve fibers and the whole nerve response shown by forward masking: An empirical model for the AP,” *J. Acoust. Soc. Am.* **73**(2), 580–591.

Don M. and Eggermont J.J. (1978):

“Analysis of the click-evoked brain stem potentials in man using high-pass noise masking,” *J. Acoust. Soc. Am.* **63**(4), 1084–1092.

Don M., Ponton C.W., Eggermont J.J., and Masuda A. (1994):

“Auditory brainstem response (ABR) peak amplitude variability reflects individual differences in cochlear response times,” *J. Acoust. Soc. Am.* **96**(6), 3476–3491.

Don M., Masuda A., Nelson R., and Brackmann D. (1997):

“Successful detection of small acoustic tumors using the stacked derived-band auditory brain stem response amplitude,” *Am. J. Otol.* **18**(5), 608–621.

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.

van der Drift J.F.C., Brocaar M.P., and van Zanten G.A. (1987):

“The relation between the pure-tone audiogram and the click auditory brainstem response threshold in cochlear hearing loss,” *Audiology* **26**(1), 1–10.

van der Drift J.F.C., Brocaar M.P., and van Zanten G.A. (1988a):

“Brainstem response audiometry. I. Its use in distinguishing between conductive and cochlear hearing loss,” *Audiology* **27**(5), 260–270.

van der Drift J.F.C., Brocaar M.P., and van Zanten G.A. (1988b):

“Brainstem response audiometry. II. Classification of hearing loss by discriminant analysis,” *Audiology* **27**(5), 271–278.

Eddins A.C. (1998):

“Spectro-temporal influence on auditory cortical evoked potential thresholds,” in *16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America* Seattle, Washington pp. 879–880.

Eggermont J.J. (1976):

“Electrocochleography,” in *Handbook of Sensory Physiology*, edited by W.D. Keidel and W.D. Neff (Springer, Berlin), volume V/3 pp. 625–705.

Eggermont J.J. and Don M. (1980):

“Analysis of the click-evoked brain stem potentials in man using high-pass noise masking. II. Effect of click intensity,” *J. Acoust. Soc. Am.* **68**(6), 1671–1675.

Elberling C. (1976):

“Modeling action potentials,” *Rev. Laryngol. Otol. Rhinol. (Bord)* **97 Suppl.**, 527–537.

Evans E.F. and Elberling C. (1982):

“Location-specific components of the gross cochlear action potential: An assessment of the validity of the high-pass masking technique by cochlear nerve fibre recording in the cat,” *Audiology* **21**(3), 204–227.

Folsom R.C. (1984):

“Frequency specificity of human auditory brainstem potentials as revealed by pure-tone masking profiles,” *J. Acoust. Soc. Am.* **75**(3), 919–924.

Folsom R.C. (1985):

“Auditory brainstem responses from human infants: Pure-tone masking profiles for clicks and filtered clicks,” *J. Acoust. Soc. Am.* **78**(2), 555–562.

Galambos R., Makeig S., and Talmachoff P.J. (1981):

“A 40-Hz auditory potential recorded from the human scalp,” *Proc. Natl. Acad. Sci. U.S.A.* **78**(4), 2643–2647.

Galbraith G.C. (1994):

“Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli,” *Electroenceph. clin. Neurophysiol.* **92**(4), 321–330.

Geisler C.D., Rhode W.S., and Kennedy D.T. (1974):

“Responses to tonal stimuli of single auditory nerve fibers and their relation to basilar membrane motion in the squirrel monkey,” *J. Neurophysiol* **37**(6), 1156–1172.

Glasberg B.R. and Moore B.C.J. (1990):

“Derivation of auditory filter shapes from notched-noise data,” *Hear. Res.* **47**(1–2), 103–138.

Goldstein M.H. and Kiang N.Y.S. (1958):

“Synchrony of neural activity in electric responses evoked by transient stimuli,” *J. Acoust. Soc. Am.* **30**(2), 107–114.

Gorga M.P. and Thornton A.R. (1989):

“The choice of stimuli for ABR measurements,” *Ear Hear.* **10**(4), 217–230.

Gorga M.P., Worthington D.W., Beauchaine K.A., and Goldgar D.E. (1985):

“Some comparisons between brain-stem response threshold, latencies and the pure tone audiogram,” *Ear Hear.* **6**(2), 105–112.

Gorga M.P., Kaminski J.R., Beauchaine K.A., and Jestead W. (1988):

“Auditory brainstem responses to tone bursts in normal hearing subjects,” *J. Speech Hear. Res.* **31**(1), 87–97.

Granzow M., Riedel H., and Kollmeier B. (2001):

“Single-sweep-based methods to improve the quality of auditory brain stem responses. Part I: Optimized linear filtering,” *Z. Audiol.* **40**(1), 32–44.

Greenwood D.D. (1990):

“A cochlear frequency position function for several species – 29 years later,” *J. Acoust. Soc. Am.* **87**(6), 2592–2605.

Griffiths S.K. and Chambers R.D. (1991):

“The amplitude modulation-following response as an audiometric tool,” *Ear Hear.* **12**(4), 235–241.

Gummer A.W. and Zenner H.P. (1996):

“Central processing of auditory information,” in *Comprehensive human physiology*, edited by R. Greger and U. Windhorst (Springer, Berlin), volume 1 pp. 227–234.

Gutschalk A., Mase R., Roth R., Ille N., Rupp A., Hähnel S., Picton T.W., and Scherg M. (1999):

“Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex,” *Clin. Neurophysiol.* **110**(5), 856–868.

Hansen P.C. (1994):

“Regularization Tools: A Matlab package for analysis and solution of discrete ill-posed problems,” *Numerical Algorithms* **6**, 1–35.

Hansen P.C. (1997):

*Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion.* (SIAM, Philadelphia).

Hashimoto I. (1982):

“Auditory evoked potentials from the human midbrain: Slow brainstem responses,” *Electroenceph. clin. Neurophysiol.* **53**(6), 652–657.

Hecox K.E., 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), 1178–1192.

Heinz M.G., Zhang X., Bruce I.C., and Carney L.H. (2001):

“Auditory nerve model for predicting performance limits of normal and impaired listeners,” *Acoustics Research Letters Online* **2**(3), 91–96.

Hoke M., Pantev C., Ansa L.M., Lütkenhöner B., and Herrmann E. (1991):

“A timesaving BERA technique for frequency-specific assessment of the auditory threshold through tone-pulse series stimulation (TOPSTIM) with simultaneous gliding high-pass noise masking (GHINOMA),” *Acta Otolaryngol. Suppl.* **482**, 45–56.

Hoth S. and Lenarz T. (1994):

*Elektrische Reaktionsaudiometrie* (Springer, Berlin).

Hou S.M. and Lipscomb D.M. (1979):

“An investigation of the auditory frequency-following responses as compared to cochlear potentials,” *Arch. Oto-Rhino-Laryngol.* **222**, 235–240.

Jacobson J.T. (1983):

“Effects of rise time and noise masking on tone pip auditory brainstem responses,” *Sem. Hear.* **4**, 363–372.

Janssen T., Steinhoff H.J., and Böhnke F. (1991):

“Zum Entstehungsmechanismus der Frequenzfolgepotentiale,” *Oto-Rhino-Laryngologia Nova* **1**, 16–25.

Jasper H.H. (1957):

“The ten twenty electrode system of the international federation,” *Electroenceph. clin. Neurophysiol.* **10**, 371–375, appendix.

Jewett D.L. (1970):

“Volume-conducted potentials in response to auditory stimuli as detected by averaging in the cat,” *Electroenceph. clin. Neurophysiol.* **28**, 609–618.

Kemp D.T. (1978):

“Stimulated acoustic emissions from within the human auditory system,” *J. Acoust. Soc. Am.* **64**(5), 1386–1391.

Kileny P. (1981):

“The frequency specificity of tone-pip evoked auditory brain stem responses,” *Ear Hear.* **2**(6), 270–275.

Klein A.J. (1983):

“Properties of the brain-stem response slow-wave component. II. Frequency specificity,” *Arch. Otolaryngol.* **109**(2), 74–78.

Klein A.J. and Mills J.H. (1981a):

“Physiological (wave I and V) and psychological tuning curves in human subjects,” *J. Acoust. Soc. Am.* **69**(3), 760–768.

Klein A.J. and Mills J.H. (1981b):

“Physiological and psychological measures from humans with temporary threshold shift,” *J. Acoust. Soc. Am.* **70**(4), 1045–1053.

Klein A.J. and Teas D.C. (1978):

“Acoustically dependent latency shifts of BSER (wave V) in man,” *J. Acoust. Soc. Am.* **63**(6), 1887–1895.

Kodera K., Yamone H., Yamada O., and Suzuki J.I. (1977):

“The effect of onset, offset and rise-decay times of tone bursts on brainstem response,” *Scand. Audiol.* **6**(4), 205–210.

Kramer S.J. (1992):

“Frequency-specific auditory brainstem responses to bone-conducted stimuli,” *Audiology* **31**(2), 61–71.

Kramer S.J. and Teas D.C. (1979):

“BSR (wave V) and N1 latencies in response to acoustic stimuli with different bandwidths,” *J. Acoust. Soc. Am.* **66**(2), 446–455.

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.

Laukli E. (1983a):

“Stimulus waveforms used in brainstem response audiometry,” *Scand. Audiol.* **12**(2), 83–89.

Laukli E. (1983b):

“High-pass and notch noise masking in suprathreshold brainstem response audiometry,” *Scand. Audiol.* **12**(2), 109–115.

Laukli E. and Mair I.W.S. (1986):

“Frequency specificity of the auditory brainstem responses. A derived band study,” *Scand. Audiol.* **15**(3), 141–146.

Laukli E., Fjermedal O., and Mair I.W.S. (1988):

“Low-frequency and auditory brainstem response threshold,” *Scand. Audiol.* **17**(3), 171–178.

Liègeois-Chauvel C., Musolino A., Badier J.M., Marquis P., and Chauvel P. (1994):

“Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components,” *EEG* **92**, 204–214.

Lütkenhöner B., Kauffmann G., Pantev C., and Ross B. (1990):

“Verbesserung der Synchronisation auditorisch evozierter Hirnstammpotentiale durch Verwendung eines die kochleären Laufzeitunterschiede kompensierenden Stimulus,” *Arch. Otolaryngol. Suppl.* II pp. 157–159.



Mackersie C., Down K.E., and Stapells D.R. (1993):

“Pure-tone masking profiles for human auditory brainstem and middle-latency responses,” *Hear. Res.* **65**(1–2), 61–68.

Melcher J.R. and Kiang N.Y.S. (1996):

“Generators of the brainstem auditory evoked potential in cat. III: Identified cell populations,” *Hear. Res.* **93**(1–2), 52–71.

Møller A.R. and Jannetta P.J. (1982):

“Evoked potentials from the inferior colliculus in man,” *Electroenceph. clin. Neurophysiol.* **53**(6), 612–620.

Møller A.R. and Jannetta P.J. (1986):

“Simultaneous surface and direct brainstem recordings of brainstem auditory evoked potentials (BAEP) in man,” in *Evoked Potentials*, edited by R.Q. Cracco and I. Bodis-Wollner (Alan R. Liss, New York), pp. 227–234.

Moore B.C. and Glasberg B.R. (1997):

“A model of loudness perception applied to cochlear hearing loss,” *Aud. Neurosci.* **3**, 289–311.

Moore J.K. (1987a):

“The human auditory brain stem: A comparative view,” *Hear. Res.* **29**(1), 1–32.

Moore J.K. (1987b):

“The human auditory brain stem as a generator of auditory evoked potentials,” *Hear. Res.* **29**(1), 33–43.

Neely S.T., Norton S.J., Gorga M.P., and Jesteadt W. (1988):

“Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli,” *J. Acoust. Soc. Am.* **83**(2), 652–656.

Neumann J., Uppenkamp S., and Kollmeier B. (1994):

“Chirp evoked otoacoustic emissions,” *Hear. Res.* **79**(1–2), 17–25.

Norton S.J. and Neely S.T. (1987):

“Tone-burst-evoked oto-acoustic emissions in normal hearing subjects,” *J. Acoust. Soc. Am.* **81**(6), 1860–1872.

Nousak J.M.K. and Stapells D.R. (1992):

“Frequency specificity of the auditory brain stem response to bone-conducted tones in infants and adults,” *Ear Hear.* **13**(2), 87–95.

Nousak J.M.K. and Stapells D.R. (1998):

“Loudness and the ABR/MLR in noise-masked normal-hearing subjects,” in *21th meeting of the Association for Research in Otolaryngology*.

Oates P.A. and Stapells D.R. (1997a):

“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.

Oates P.A. and Stapells D.R. (1997b):

“Frequency specificity of the human auditory brainstem and middle-latency responses to brief tones. II. Derived response analysis,” *J. Acoust. Soc. Am.* **102**(6), 3609–3619.

Pantev C., Lagidze S., Pantev M., and Kevanishvili Z. (1985):

“Frequency-specific contributions to the auditory brain stem response derived by means of pure-tone masking,” *Audiology* **24**(4), 275–287.

Patterson R.D., Robinson K., Holdsworth J., McKeown D., Zhang C., and Allerhand M. (1992):

“Complex sounds and auditory images,” in *Auditory physiology and perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 429–446.

Patterson R.D., Allerhand M., and Giguère C. (1995):

“Time-domain modelling of peripheral auditory processing: A modular architecture and a software platform,” *J. Acoust. Soc. Am.* **98**(4), 1890–1894.

Picton T.W., Hillyard S.A., Krausz H.I., and Galambos R. (1974):

“Human auditory evoked potentials. I: Evaluation of components,” *Electroenceph. clin. Neurophysiol.* **36**, 179–190.

Picton T.W., Ouellette J., Hamel G., and Durieux-Smith A. (1979):

“Brainstem evoked potentials to tonepips in notched noise,” *J. Otolaryngol.* **8**(4), 289–314.

Pratt H. and Bleich N. (1982):

“Auditory brainstem potentials evoked by clicks in notch-filtered masking noise,” *Electroenceph. clin. Neurophysiol.* **53**(4), 417–426.

Purdy S.C., Houghton J.M., Keith W.J., and Greville K.A. (1989):

“Frequency-specific auditory brainstem responses. Effective masking levels and relationship to behavioural thresholds in normal hearing adults,” *Audiology* **28**(2), 82–91.

Rhode W.S. (1971):

“Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique,” *J. Acoust. Soc. Am.* **49**(4), 1218–1231.

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), 62–65.

Robles L., Ruggero M.A., and Rich N.C. (1986):

“Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases,” *J. Acoust. Soc. Am.* **80**(5), 1364–1374.

Rose J.E., Hind J.E., Anderson D.J., and Brugge J.F. (1971):

“Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey,” *J. Neurophysiol.* **34**(4), 685–699.

Ruggero M.A. (1992):

“Responses to sound of the basilar membrane of the mammalian cochlea,” *Curr. Opin. Neurobiol.* **2**(4), 449–456.

Ruggero M.A. and Rich N.C. (1983):

“Chinchilla auditory nerve responses to low-frequency tones,” *J. Acoust. Soc. Am.* **73**(6), 2096–2108.

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.

Rupp A., Uppenkamp S., Gutschalk A., Patterson R.D., Dau T., and Scherg M. (2002):

“On the representation of peripheral neural activity in primary auditory cortex,” *Hear. Res.* (in press).

Russell I.J. and Sellick P.M. (1978):

“Intracellular studies of hair cells in the mammalian cochlea,” *J. Physiol.* **284**, 261–290.

Scherg M. and von Cramon D. (1985):

“A new interpretation of the generators of BAEP waves I-V: Results of a spatio-temporal dipole model,” *Electroenceph. clin. Neurophysiol.* **62**(4), 290–299.

Scherg M. and von Cramon D. (1990):

“Dipole source potentials of the auditory cortex in normal subjects and in patients with temporal lobe lesions,” in *Advances in Audiology*, edited by M. Hoke (Karger, Basel), pp. 165–192.

Scherg M. and Volk S.A. (1983):

“Frequency specificity of simultaneously recorded early and middle-latency auditory evoked potentials,” *Electroenceph. clin. Neurophysiol.* **56**(5), 443–452.

Scherg M., Hari R., and Hämäläinen M. (1989):

“Frequency-specific sources of the auditory N19-P30-P50 response detected by multiple source analysis of evoked magnetic fields and potentials,” in *Advances in Biomagnetics*, edited by S.J. Williamson, M. Hoke, and G. Sroink (Plenum Press, New York), pp. 97–100.

Sellick P.M., Patuzzi R., and Johnstone B.M. (1982):

“Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique,” *J. Acoust. Soc. Am.* **72**(1), 131–141.

Shera C.A. and Guinan J.J. (2000):

“Frequency dependence of stimulus-frequency-emission phase: Implications for cochlear mechanics,” in *Recent Developments in Auditory Mechanics*, edited by H. Wada,

T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 381–387.

Shera C.A. and Zweig G. (1993):

“Order from the chaos: Resolving the paradox of periodicity in evoked otoacoustic emission,” in *Biophysics of Hair-Cell Sensory Systems*, edited by H. Duifhuis, J.W. Horst, P. van Dijk, and S.M. van Netten (World Scientific, Singapore), pp. 54–63.

Shera C.A., Guinan J.J., and Oxenham A.J. (2002):

“Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements,” *Proc. Natl. Acad. Sci. USA* **99**(5), 3318–3323.

Shore S.E. and Nuttall A.L. (1985):

“High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts,” *J. Acoust. Soc. Am.* **78**(4), 1286–1295.

Shore S.E., Clopton B.M., and Au Y.N. (1987):

“Unit responses in ventral cochlear nucleus reflect cochlear coding of rapid frequency sweeps,” *J. Acoust. Soc. Am.* **82**(2), 471–478.

Sohmer H. and Kinarti R. (1984):

“Survey of attempts to use auditory evoked potentials to obtain an audiogram,” *Br. J. Audiol.* **18**(4), 327–244.

Stapells D.R. and Picton T.W. (1981):

“Technical aspects of brainstem evoked potential audiometry using tones,” *Ear Hear.* **2**(1), 20–29.

Stapells D.R., Picton P.E., Perez-Abalo M.C., Read D., and Durieux-Smith A. (1990):

“Frequency specificity in evoked potential audiometry,” in *The Auditory Brainstem Response*, edited by J.T. Jacobson (College-Hill, San Diego), pp. 147–177.

Suzuki T., Hirai Y., and Hiriuchi K. (1977):

“Auditory brain stem responses to pure tone stimuli,” *Scand. Audiol.* **6**(1), 51–56.

Terkildsen K., Osterhammel P., and in't Velt F.H. (1975):

“Farfield electrocochleography. Frequency selectivity of the response,” *Scand. Audiol.* **4**, 167–172.

Tikhonov A.N. (1963):

“Solution of incorrectly formulated problems and the regularization method,” *Soviet. Math. Doct.* **4**, 1035–1038.

Tsuchitani C. (1983):

“Physiology of the auditory system,” in *Bases of Auditory Brain-Stem Evoked Responses*, edited by E.J. Moore (Grune and Stratton, New York), .

Verhey J. and Kollmeier B. (1998):

“Messungen zur zeitabhngigen Lautheitssummation,” in *Fortschritte der Akustik - DAGA 98*, edited by A. Sill (Deutsche Gesellschaft für Akustik e.V., Zürich), pp. 482–483.

Weber B.A. (1987):

“Assessing low-frequency hearing using auditory evoked potentials,” *Ear Hear.* **8**(Suppl. 4), 49–54.

Wegner O. and Dau T. (2002):

“Frequency specificity of chirp-evoked auditory brainstem responses,” *J. Acoust. Soc. Am.* **111**(3), 1318–1329.

Wegner O., Dau T., Verhey J., and Kollmeier B. (1998):

“Untersuchungen zu Lautheits-Korrelaten in akustisch evozierten Potentialen,” in *Fortschritte der Akustik - DAGA 98*, edited by A. Sill (Deutsche Gesellschaft für Akustik e.V., Zürich), pp. 308–309.

Wegner O., Dau T., and Kollmeier B. (1999):

“On the relationship between auditory evoked potentials and psychophysical loudness,” in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific, Singapore), pp. 59–62.

Wu C.Y. and Stapells D.R. (1994):

“Pure-tone masking profiles for human auditory brainstem and middle latency responses to 500-Hz tones,” *Hear. Res.* **78**(2), 169–174.

Zhang X., Heinz M.G., Bruce I.C., and Carney L.H. (2001):

“A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression,” *J. Acoust. Soc. Am.* **109**(2), 648–670.

Zweig G. (1976):

“Basilar membrane motion,” *Cold Spring Harb. Symp. Quant. Biol.* **40**, 619–633.

Zweig G. and Shera C.A. (1995):

“The origin of periodicity in the spectrum of otoacoustic emissions,” *J. Acoust. Soc. Am.* **98**(4), 2018–2047.

Zwicker E. and Fastl H. (1990):

*Psychoacoustics. Facts and Models* (Springer-Verlag, Berlin).





# Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Oldenburg, den 17. Dezember 2002

Oliver Wegner



# Danksagung

An dieser Stelle möchte ich den vielen Personen danken, die mich während und bei meiner Doktorarbeit unterstützt haben. Sicherlich kann ich nicht alle aufführen, aber einige seien hier genannt:

Herrn Prof. Dr. Dr. Birger Kollmeier danke ich dafür, dass er mir die Anfertigung dieser Arbeit in einer Arbeitsgruppe mit ausgezeichneten Arbeitsbedingungen überhaupt erst ermöglicht hat. Insbesondere die Freiheiten, die er mir während der Arbeit gelassen hat und die Diskussionen haben wesentlich zum Gelingen beigetragen.

Herrn Prof. Dr. Volker Mellert danke ich für sein Interesse und für die Übernahme des Korreferats.

Mein besonderer Dank gilt Dr. Torsten Dau für seine hervorragende Betreuung in allen Phasen der Doktorarbeit, insbesondere für die intensiven Diskussionen, Korrekturen und Aufmunterungen in der Schlußphase der Arbeit.

Für das Korrekturlesen der verschiedenen Teile und Versionen der Arbeit möchte ich Torsten Dau, Helmut Riedel, Stefan Uppenkamp und Sandra Fobel danken.

Allen Mitgliedern der Arbeitsgruppe „Medizinische Physik“ danke ich für die angenehme Arbeitsatmosphäre und die vielen anregenden Gespräche, deren Thema nicht nur die Arbeit waren. An dieser Stelle sei auch den vielen geduligen Versuchspersonen gedankt, die zum großen Teil aus der Arbeitsgruppe kommen; denn es sollte nicht unterschätzt werden, wie anstrengend das stundenlange „Nichtstun“ während der Messungen sein kann. Anita Gorges möchte ich dafür danken, dass sie viele der Messungen für mich durchgeführt hat; denn auch das Überwachen der Messungen ist nicht immer wirklich spannend.

Meinen verschiedenen „Mitbewohnern“ möchte ich für die nie langweilige Arbeitsatmosphäre danken, in der insbesondere immer dann Ablenkung möglich war, wenn wieder einmal ein „Denkstau“ vorlag.

Angelehnt an das englische „last but not least“ möchte ich zu guter Letzt noch Sandra Fobel für die vielfältige Unterstützung danken die ich von ihr bekommen habe. Insbesondere in der Endphase der Arbeit war sie eine wichtige Stütze für mich.

Diese Arbeit wurde finanziell unterstützt von der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Graduiertenkollegs „Psychoakustik“ und des Sonderforschungsbereichs „Neuronale Grundlagen kognitiver Leistungen“ (SFB 517).



# Lebenslauf

Am 20. Oktober 1971 wurde ich, Oliver Wegner, als zweites Kind von Ursula Wegner, geb. Hinrichs, und Gerhard Klaus Wegner in Oldenburg mit deutscher Staatsangehörigkeit geboren.

Im Schuljahr 1978/79 wurde ich in die Grundschule Hahn-Lehmden eingeschult. Nach Beendigung der Grundschule begann ich im August 1982 auf der Orientierungsstufe der Kooperativen Gesamtschule Rastede die 5. Klasse. Im Juli 1984 ging ich von der Orientierungsstufe mit der Empfehlung an das Gymnasium ab.

Ab August 1984 besuchte ich den Gymnasialzweig der Kooperativen Gesamtschule Rastede. Dort erlangte ich im August 1988 den erweiterten Sekundarabschluß I, so daß ich im folgendem Schuljahr die gymnasiale Oberstufe der Kooperativen Gesamtschule besuchen konnte. Meine naturwissenschaftlich ausgerichteten Interessen fanden in der Wahl der Leistungskurse Mathematik und Chemie ihren Ausdruck. Am 5. Juni 1991 beendete ich meine Schullaufbahn mit dem Abitur.

Im Wintersemester 1991/92 begann ich dann mit dem Studium der Physik an der Carl-von-Ossietzky-Universität Oldenburg. Mein Vordiplom bestand ich am 31. August 1993 mit der Note „gut“.

Im Rahmen des Hauptstudiums fertigte ich im April und Mai 1995 in der Arbeitsgruppe „Angewandte Optik/Kohärente Optik“ von Herrn Professor Hinsch eine Studienarbeit mit dem Titel „Untersuchung der Eigenschaften eines Flüssigkristalls für die optische Phasenschiebung“ an.

Im August 1996 begann ich meine Diplomarbeit in der Arbeitsgruppe „Medizinische Physik“ von Herrn Professor Kollmeier. Die Arbeit wurde von Herrn Professor Kollmeier und Herrn Doktor Dau betreut. Der Titel der Arbeit lautete „Zusammenhang zwischen Psychoakustik und Hirnstammpotentialen“. Das Diplom habe ich am 13. Juni 1997 mit der Note „sehr gut“ bestanden.

Nach dem Diplom arbeitete ich fünf Monate (von August bis Dezember 1997) als wissenschaftlicher Angestellter im BMBF-Projekt „Hörgeräte-Algorithmen“ an der Carl-von-Ossietzky Universität Oldenburg. Im Oktober begann ich unter Leitung von Herrn Professor Kollmeier mit meiner Promotion am Fachbereich Physik.

Vom Februar 1998 bis zum Juni 1999 war ich Stipendiat am von der DFG geförderten Graduiertenkolleg „Psychoakustik“ an der Carl-von-Ossietzky-Universität Oldenburg. Seit Juli 1999 arbeite ich als wissenschaftlicher Angestellter im Sonderforschungsbereich 517 („Neuronale Grundlagen kognitiver Leistungen“).