Mathematics for Industry 20

Yoshiharu Soeta
Yoichi Ando

Neurally Based Measurement and Evaluation of Environmental Noise

Springer

Mathematics for Industry

Volume 20

Aims & Scope

The meaning of "Mathematics for Industry" (sometimes abbreviated as MI or MfI) is different from that of "Mathematics in Industry" (or of "Industrial Mathematics"). The latter is restrictive: it tends to be identified with the actual mathematics that specifically arises in the daily management and operation of manufacturing. The former, however, denotes a new research field in mathematics that may serve as a foundation for creating future technologies. This concept was born from the integration and reorganization of pure and applied mathematics in the present day into a fluid and versatile form capable of stimulating awareness of the importance of mathematics in industry, as well as responding to the needs of industrial technologies. The history of this integration and reorganization indicates that this basic idea will someday find increasing utility. Mathematics can be a key technology in modern society.

The series aims to promote this trend by (1) providing comprehensive content on applications of mathematics, especially to industry technologies via various types of scientific research, (2) introducing basic, useful, necessary and crucial knowledge for several applications through concrete subjects, and (3) introducing new research results and developments for applications of mathematics in the real world. These points may provide the basis for opening a new mathematics-oriented technological world and even new research fields of mathematics.

More information about this series at http://www.springer.com/series/13254

Yoshiharu Soeta · Yoichi Ando

Neurally Based Measurement and Evaluation of Environmental Noise

Springer

Yoshiharu Soeta
National Institute of Advanced Industrial Science and Technology
Osaka
Japan

Yoichi Ando
Kobe University
Kobe
Japan

ISSN 2198-350X ISSN 2198-3518 (electronic)
Mathematics for Industry
ISBN 978-4-431-55431-8 ISBN 978-4-431-55432-5 (eBook)
DOI 10.1007/978-4-431-55432-5

Library of Congress Control Number: 2015937752

Springer Tokyo Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer Japan KK is part of Springer Science+Business Media (www.springer.com)

Preface

Environmental noise is a big problem for human beings. Initially, noise level was the biggest problem. It has caused many bad effects, such as headache, discomfort, hearing loss, sleep disturbance. Long-term environmental noises also have deep effects and are integrated in the brain and body without any conscious awareness of them, as is discussed in Chap. 9. Particularly, the effects of the developments of unborn babies and the specialization of the cerebral hemispheres in growing children are serious. Numerous new technologies, such as noise insulation and active noise cancelling, have been developed and have contributed to the reduction of noise levels. But people feel annoyed even when the noise level is quite low because of the qualitative aspects of noises. In addition, the method of measuring and evaluating noise is still premature because the model of the auditory-brain system has ended at the peripheral level. Therefore, the development of measuring and evaluating the quality of noise based on our brain function is becoming more important for human beings.

This book deals with the methods of measurement and evaluation of environmental noise based on an auditory neural and brain-oriented model. The model consists of the autocorrelation function (ACF) and the interaural cross-correlation function (IACF) mechanisms for signals arriving at the two ear entrances. This model was based on neural evidence. First, we focused on the human auditory system, which was investigated mainly through the human brain and psychological response, such as by electroencephalography (EEG), magnetoencephalography (MEG), loudness, and annoyance, because the features of the human auditory system have to be taken into account in evaluating the quality of noises. The results suggest that the human auditory system has ACF and IACF mechanisms and the factors extracted from those mechanisms are useful as cues for temporal and spatial sensation of sounds.

It is hoped that the survey presented here will encourage researchers, students, and engineers in a wide range of fields, such as the automotive industry, the aerospace industry, mechanical engineering, railways, electronics industries, soundscape, architecture, and acoustics.

This book largely serves as a record of the research carried out at the Ando Laboratory, Graduate School of Science and Technology, Kobe University, between 1969 and 2009, even after the authors' graduation or retirement; as well as the studies conducted at the Living Informatics Research Group, National Institute of Advanced Industrial Science and Technology (AIST) between 2002 and 2015. The authors thank Dr. Shin-ichi Sato, Dr. Hiroyuki Sakai, Dr. Ryota Shimokura, Dr. Kenji Fujii, Dr. Seiji Nakagawa, and Dr. Mitsuo Tonoike for their collaboration and help.

March 2015

Yoshiharu Soeta
Yoichi Ando

Contents

1 Introduction 1
References. 2

2 Signal Processing Model of Human Auditory System 5
2.1 Human Hearing System. 5
2.2 Neural Evidences of the Autocorrelation Model in the Auditory Pathways. 9
2.3 Neural Evidences of the Interaural Cross-Correlation Model in the Auditory Pathways. 17
2.4 Signal Processing Model of Human Auditory System 29
2.5 Brain Response in Relation to Loudness 32
2.6 Brain Response Corresponding to Annoyance. 38
References. 44

3 Noise Measurement Method Based on the Model 51
3.1 Correlation Function 51
3.2 Temporal Factors Extracted from the ACF. 53
3.3 Temporal Window of the ACF Processing 59
3.4 Spatial Factors Extracted from the IACF 61
3.5 Auditory Temporal Window for the IACF Processing 62
3.6 System of Noise Measurement 63
References. 66

4 Temporal Primary Sensations of Noise 69
4.1 Formulation of Temporal and Spatial Sensations. 69
4.2 Loudness 71
4.2.1 Loudness of Sharply Filtered Noise. 71
4.2.2 Loudness of Complex Noise 74
4.2.3 Loudness of Iterated Rippled Noise. 79

4.3 Pitch . . . 88
4.3.1 Pitches of Complex Tones and "Complex Noise" . . . 88
4.3.2 Pitch of "Complex Noise" . . . 93
4.3.3 Frequency Limits of the ACF Model . . . 95
4.4 Timbre . . . 97
4.5 Duration . . . 99
References . . . 101

5 Spatial Primary Sensations of Noise . . . 105
5.1 Localization of Noise Source . . . 105
5.2 Subjective Diffuseness . . . 107
5.3 Apparent Source Width (ASW) . . . 112
5.3.1 ASW of the Band-Pass-Filtered Noise . . . 112
5.3.2 ASW of "Complex Noise" . . . 116
References . . . 118

6 Noise Measurements . . . 121
6.1 Aircraft Noise . . . 122
6.2 Road Traffic Noise . . . 128
6.3 Flushing Toilet Noise . . . 132
6.4 Railway Noise in a Train Station . . . 137
6.5 Railway Noise in a Train Car . . . 144
6.5.1 Effect of Train Type and Noise Caused by Wheel–Rail Interaction . . . 144
6.5.2 Effects of External Environments . . . 151
6.6 Floor Impact Sound . . . 155
6.7 Heating, Ventilating, and Air Conditioning (HVAC) Noise . . . 162
References . . . 162

7 Annoyance of Noise . . . 167
7.1 Annoyance of Noise with a Pure Tone Component . . . 167
7.2 Annoyance of Band-Pass-Filtered Noise . . . 173
7.3 Annoyance of Noise in Relation to the Spatial Factors . . . 179
7.3.1 Effect of IACC Fluctuation . . . 179
7.3.2 Effect of τ_{IACC} Fluctuation . . . 181
7.4 Annoyance of Road Traffic Noise . . . 185
7.5 Annoyance of Noise in a Train Station . . . 188
7.6 Annoyance of Noise in a Train Car . . . 191
7.7 Annoyance of Floor Impact Sound . . . 194
7.8 Annoyance of Heating, Ventilation, and Air Conditioning (HVAC) Noise . . . 196
7.9 A General Equation for Annoyance . . . 198
References . . . 200

8 Short-Term Effects of Noise 203
8.1 Speech Disturbance by Different Directional Noise 203
8.2 Sleep Disturbance by Upstairs Toilet Noise Changing Spatial Factors 208
8.3 Model of Duration Experience Due to Noise 210
References. 212

9 Long-Term Effects of Noise. 215
9.1 Human Placental Lactogen (HPL) in Maternal Plasma. 216
9.2 Birthweight 218
9.3 Reactions in Sleeping Babies 220
9.3.1 Reaction of Babies on PLG and EEG 220
9.3.2 Remarks of ACF Factors of the Noise and Music 222
9.4 Development of Specialization of Cerebral Hemispheres 224
9.5 Development of Height 228
References. 230

10 Application to Sound Design 233
10.1 Subjective Diffuseness of Music Signals Convolved with Binaural Impulse Responses 233
10.1.1 Relationship Between $IACC_{IR}$ and $IACC_{SR}$ 235
10.1.2 PsychoAcoustic Experiment 239
10.2 Listening Level of Music Through Headphones in Train Car Noise Environments 243
10.2.1 Experimental Methods. 244
10.2.2 Effects of Noise Source 248
10.2.3 Effects of Music Source 250
10.3 Subject Preference for Birdsongs 253
10.4 Urban Soundscape Design 256
References. 257

Index 261

Abbreviations

ABR	Auditory brain stem response
AC	Auditory cortex
ACF	Autocorrelation function
AEF	Auditory evoked magnetic field
AEP	Auditory evoked potential
AGC	Automatic gain control
ANOVA	Analysis of variance
ASW	Apparent source width
BC	Bone chain
BIR	Binaural impulse response
BIRs	Binaural impulse responses
BM	Basilar membrane
BMM	Boring machine method
CCF	Cross-correlation function
CCM	Cut-and-cover method
CN	Cochlear nucleus
DL	Difference limen
EC	External canal
ECD	Equivalent current dipole
ED	Eardrum
EEG	Electroencephalography
EOR	Energy onset response
EPNL	Effective perceived noise level
FFR	Frequency following response
GME	Ear canal sound pressure to cochlear vestibule pressure gain
HC	Hair cell
HPL	Human placental lactogen
HRTF	Head-related transfer function
HVAC	Heating, ventilation, and air conditioning
IACC	Interaural cross-correlation coefficient
$IACC_{IR}$	IACC of impulse response

$IACC_{SR}$	IACC of other signal response through a hall
IACF	Interaural cross-correlation function
IC	Inferior colliculus
IRN	Iterated rippled noise
ITD	Interaural time delay
JND	Just-noticeable difference
LEV	Listener envelopment
LLN	Lateral lemniscus nucleus
LSO	Lateral superior olive
MEG	Magnetoencephalography
MGB	Medial denticulate body
MLS	Maximum length sequence
MNTB	Medial nucleus of the trapezoid body
MSO	Medial superior olive
NACF	Normalized autocorrelation function
NATM	New Austrian tunneling method
NC	Noise criterion
NCB	Balanced noise criterion
NI	Nonidentification
PLG	Electroplethysmography
PNC	Preferred noise criterion
PNL	Perceived noise level
PNLT	Tone-corrected perceived noise level
POR	Pitch onset response
PSDs	Platform screen doors
PSE	Point of subjective equality
RC	Room criteria
RMS	Root-mean-squares
RNC	Room noise criterion
SOC	Superior olivary complex
SV	Scale values
SVTF	Ear canal sound pressure to stapes footplate velocity transfer function
TNIS	Train noise in a station
WECPNL	Weighted equivalent continuous perceived noise level

Chapter 1
Introduction

Abstract First, background and issues of this book and related literatures are described. Second, our solution to the issues based on human auditory system is briefly explained.

Keywords Auditory system · Autocorrelation · Interaural cross-correlation

For the evaluation of environmental noise, a large number of noise indices have been proposed by many researchers. A-weighted equivalent continuous sound pressure level (SPL), L_{Aeq}, is probably the most widespread noise index. It accounts for the magnitude of a noise and the sensitivity at different frequencies. It is simple to measure and correlate well with many psychological response to noise (e.g., Namba and Kuwano 1984; Kuwano et al. 1989; Ayr et al. 2003).

Noise criterion (NC) curves were proposed for rating indoor noise, noise from air-conditioning equipment, and so on. The method consists of a set of criteria curves extending from 63 to 8000 Hz and a tangency rating procedure. The criteria curves define the limits of octave band spectra that must not be exceeded to meet occupant acceptance in certain spaces. NC curves apply to the total noise in a room, with normal office activities in progress and all mechanical systems operating (Beranek 1956). Preferred noise criterion (PNC) curves are modified versions of the original NC curves. They are lower in both the low and high frequencies than originally specified in the NC curves (Beranek 1971). Noise rating (NR) curves are based on similar assumptions. Balanced noise criterion (NCB) curves are a further improvement of NC and PNC curves (Beranek 1989). Room criteria (RC) curves were derived by Blazier (1981) for application to the acoustical design and rating of heating, ventilation, and air conditioning (HVAC) systems. The RC method has been revised to the RC Mark II method (Blazier 1995, 1997). This latter method is more complicated to use but provides more thorough information about the character of the noise. In order to reach a technical compromise between NCB and RC curves, a set of curves named room noise criterion (RNC) was proposed

Y. Soeta and Y. Ando, *Neurally Based Measurement and Evaluation of Environmental Noise*, Mathematics for Industry 20,
DOI 10.1007/978-4-431-55432-5_1

(Shomer 2000; Shomer and Bradley 2000). It can be used with a tangency method for determining an RNC rating of a room sound spectrum, but it allows to evaluate temporal variations in low frequency sound as well.

However, these indices are determined based on SPL, that is, quantitative aspects of a noise, and frequency characteristics. For evaluation of noise, qualitative aspects of a noise are important because people may feel annoyed due to the aspects of sound quality even when the SPL of a noise is only about 35 dBA (Kitamura et al. 2002). The aspects can be formulated by the factors extracted from the autocorrelation function (ACF) and the interaural cross-correlation function (IACF) of noises arriving at two ear entrances.

This book deals with indices extracted from the ACF and IACF for the evaluation of sound quality although basic psychoacoustic indices such as loudness, sharpness, roughness, and fluctuation strength have widely used as sound quality indices (Zwicker and Fastl 1999). One rationale is that the perception of most auditory quality is based on information embedded in the timing of spikes, that is, temporal correlation representations arise from spike timing patters in the auditory nerve (Yin et al. 1987; Ando et al. 1991; Cariani and Delgutte 1996a, b; Saberi et al. 1998). Another rationale is that ACF and IACF factors describe temporal primary sensations, such as pitch, loudness, and timbre, and spatial primary sensations, such as localization in the horizontal plane, apparent source width, and subjective diffuseness, respectively (Ando et al. 1999; Ando 2001; Ando and Cariani 2009).

In Chap. 2, main and basic idea of this book, that is, signal processing model based on human auditory system, is described. In Chap. 3, concept of correlation and definitions of temporal and spatial factors obtained from ACF and IACF are explained. Chapters 2 and 3 are fundamental frameworks of this book. In Chaps. 4 and 5, temporal and spatial sensations, such as loudness, pitch, and apparent source width, are formulated based on the ACF and IACF factors. In Chap. 6, some examples of noise measurements based on the ACF and IACF factors are explained. In Chap. 7, subjective annoyance to some noises is described by the ACF and IACF factors. In Chaps. 8 and 9, short-term and long-term effects of noise on human body are discussed. In Chap. 10, some applications of the model to sound design are exemplified.

References

Ando Y (2001) A theory of primary sensations and spatial sensations measuring environmental noise. J Sound Vib 241:3–18

Ando Y, Cariani P (2009) Auditory and visual sensations. Springer, New York

Ando Y, Yamamoto K, Nagamastu H, Kang SH (1991) Auditory brainstem response (ABR) in relation to the horizontal angle of sound incidence. Acoust Lett 15:57–64

Ando Y, Sato S, Sakai H (1999) Fundamental subjective attributes of sound fields based on the model of auditory brain system. In: Sendra JJ (ed) Computational acoustics in architecture. WIT Press, Southampton, pp 63–99

Ayr U, Cirillo E, Fato I, Martellotta F (2003) A new approach to assessing the performance of noise indices in buildings. Appl Acoust 64:129–145

Beranek LL (1956) Criteria for office quieting based on questionnaire rating studies. J Acoust Soc Am 28:833–852

Beranek LL (1971) Noise and vibration control McGraw-Hill, New York

Beranek LL (1989) Balanced noise-criterion (NCB) curves. J Acoust Soc Am 86:650–664

Blazier WE (1981) Revised noise criterion for application in the acoustical design and rating of HVAC systems. Noise Control Eng J 162:64–73

Blazier WE (1995) Sound quality consideration in rating noise from heating, ventilating and air-conditioning (HVAC) systems in buildings. Noise Control Eng J 43:53–63

Blazier WE (1997) RC Mark II: A refined procedure for rating the noise of heating, ventilating, and airconditioning (HVAC) systems in buildings. Noise Control Eng J 45:243–250

Cariani PA, Delgutte B (1996a) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J Neurophysiol 76:1698–1716

Cariani PA, Delgutte B (1996b) Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. J Neurophysiol 76:1717–1734

Kitamura T, Sato S, Shimokura R, Ando Y (2002) Measurement of temporal and spatial factors of a flushing toilet noise in a downstairs bedroom. J Temporal Des Arch Environ 2:13–19

Kuwano S, Namba S, Hajime Miura H (1989) Advantages and disadvantages of A-weighted sound pressure level in relation to subjective impression of environmental noise. Noise Control Eng J 33:107–115

Namba S, Kuwano S (1984) Psychological study on Leq as a measure of loud-ness of various kinds of noises. J Acoust Soc Jpn (E) 5:135–148

Saberi K, Takahashi Y, Konishi M, Albeck Y, Arthur BJ, Farahbod H (1998) Effects of interaural decorrelation on neural and behavioral detection of spatial cues. Neuron 21:789–798

Shomer PD (2000) Proposed revisions to room noise criteria. Noise Control Eng J 48:85–96

Shomer PD, Bradley JS (2000) A test of proposed revisions to room noise criteria curves. Noise Control Eng J 48:124–129

Yin TC, Chan JCK, Carney LH (1987) Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. J Neurophysiol 58:562–583

Zwicker E, Fastl H (1999) Psychoacoustics: facts and models. Springer, Berlin

Chapter 2
Signal Processing Model of Human Auditory System

Abstract To evaluate environmental noise, we need to use methods based on functioning of our auditory system. In this chapter, basics of human auditory system are described. First, the ear sensitivity of the human ear from a sound source to the auditory system consisting of the external canal, eardrum, bone chain with oval window, auditory nerve, cochlear nucleus, thalamus, brain stem, and cortex, and signal processing model for evaluation of environmental noise are described. Second, brain responses in relation to basic perception such as loudness and annoyance are described.

Keywords Auditory system · Autocorrelation model · Interaural cross-correlation model · Auditory brain stem response · Auditory evoked magnetic fields · Loudness · Annoyance

2.1 Human Hearing System

The main function of hearing system is to get information about the outside, which is carried by pressure variations in the air, that is, sound wave. Sound waves are generated by the movement or vibration of an object, that is, sound source. As the vibrating object moves out and in, the nearby air molecules create a slight increase and decrease in pressure, called condensation and rarefaction, respectively. From the pressure variations, we perceive what the sound source is and where it comes from.

We perceive a sound wave, which is a continual time series signal, by the ears. We also perceive three-dimensional acoustic space by the ears, mainly because the head-related transfer function (HRTF) between a point of a sound source and the two ear entrances has directional characteristics from the shapes of the head and the pinnae. The pinnae significantly modify the incoming sound, particularly at high frequencies, and this is important in our ability for sound localization (Gardner and Gardner 1973; Butler and Belundiuk 1977). HRTF includes the interaural time and interaural level differences.

Y. Soeta and Y. Ando, *Neurally Based Measurement and Evaluation of Environmental Noise*, Mathematics for Industry 20,
DOI 10.1007/978-4-431-55432-5_2

Figure 2.1 shows the example of amplitude of the HRTF, $H(\xi, \eta, \omega)$, as a parameter of angle of incidence ξ (Mehrgardt and Mellert 1977). The angles ξ and η indicates azimuth and elevation angle, respectively. The angle $\xi = 0°$ corresponds to the frontal direction and $\xi = 90°$ corresponds to the lateral direction toward the side of the ear being examined.

After a sound wave arrives nearby, it passes through the peripheral auditory system, the outer ear, middle ear, and inner ear. The outer ear is the external part of the auditory system, including the pinnae and the ear canal. Sound travels down the ear canal and causes the eardrum, or tympanic membrane, to vibrate. Because of the resonance of the outer ear, we are more sensitive to sound frequencies between 1000 and 6000 Hz. The transfer function of the ear canal, $E(\omega)$, is shown in Fig. 2.2 (Wiener and Ross 1946; Shaw 1974; Mehrgardt and Mellert 1977).

The middle ear is the air-filled space between the eardrum and the cochlea that contains the ossicles. The acoustic vibrations on the eardrum are transmitted through the middle ear by three small bones, malleus, incus, and stapes, to the oval window of the cochlea. The middle ear acts as an impedance-matching device or transformer that improves sound transmission and reduces the amount of reflected sound. This is accomplished mainly by the differences in effective areas of eardrum and the oval window and to a small extent by the lever action of the ossicles. Transmission of sound through the middle ear, $C(\omega)$, is most efficient at frequencies between 500 and 4000 Hz as shown in Fig. 2.3 (Puria et al. 1997; Aibara et al. 2001).

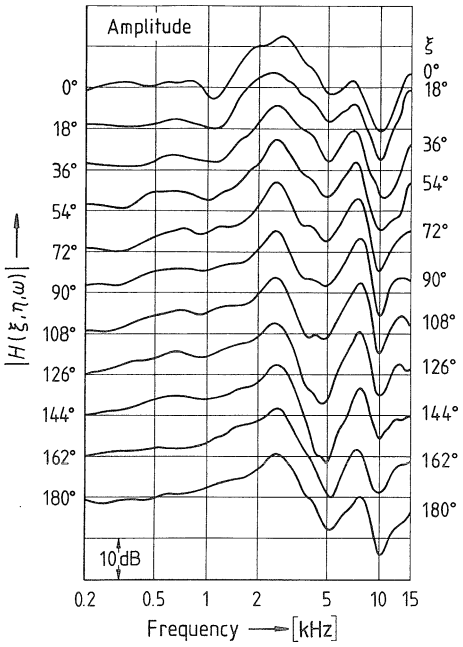


Fig. 2.1 Transfer functions (amplitude) from a free field to the ear canal entrance as a parameters of angle of incidence ξ (Mehrgardt and Mellert 1977)

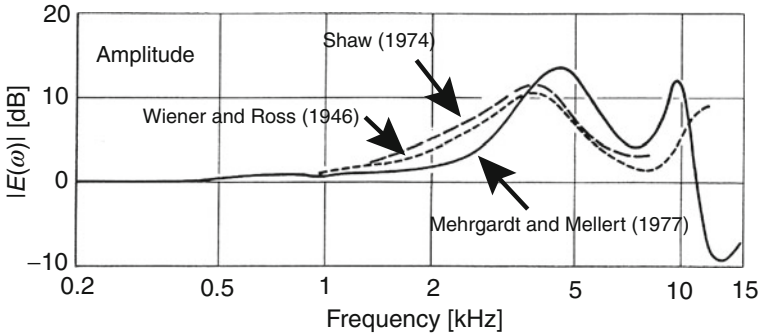
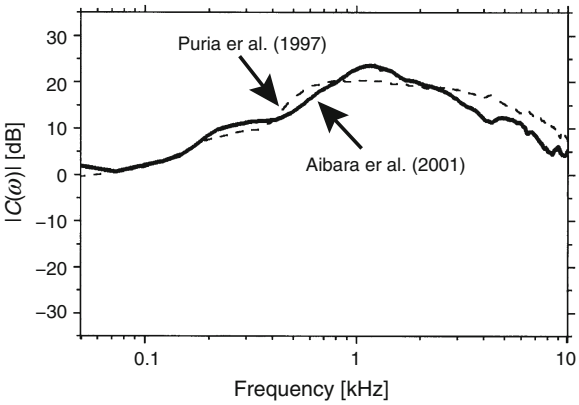


Fig. 2.2 Transfer functions (amplitude) of the ear canal (Mehrgardt and Mellert 1977)

Fig. 2.3 Transfer functions (amplitude) of the middle ear (Aibara et al. 2001)



For the usual sound field, the transfer function between a sound source located in front of the listener and the cochlea, $S(\omega)$, may be represented by

$$S(\omega) = H(\xi, \eta, \omega)E(\omega)C(\omega). \tag{2.1}$$

The values are plotted in Fig. 2.4 (Ando 1998). The pattern of the transfer function agrees with the ear sensitivity for people with normal hearing estimated from equal-loudness-level contours at 40 phon (ISO 226:2003).

The inner ear is the part of the ear that is filled with fluid, including the cochlea and semicircular canals. Sound enters the cochlea through the oval window covered by a membrane. When the oval window moves due to the pressure from the stapes, Reissner's membrane and the basilar membrane are pushed down, and the round window moves out. It follows that vibration of the stapes leads to vibration of the basilar membrane. The basilar membrane separates out the frequency components of a sound. At the base of the cochlea, near the oval window, the basilar membrane is narrow and stiff and is sensitive to high frequencies. At the apex of the cochlea,

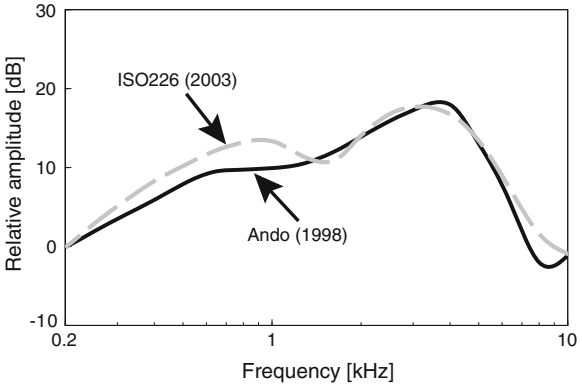


Fig. 2.4 Sensitivity of the human ear to a sound source in front of the listeners estimated from transformation characteristics between the sound source and the cochlea, $S(\omega) = H(\omega)E(\omega)C(\omega)$, (Ando 1998) and equal-loudness-level contours at 40 phon (ISO 226)

the other end of the membrane is wide and loose and is sensitive to low frequencies. The basilar membrane behaves as a band of overlapping band-pass filters, which is called auditory filters.

The mechanical vibrations of the basilar membrane are converted into electrical activity in the auditory nerve. This task is accomplished by the inner hair cells. Vibration of the basilar membrane causes a displacement of a stereochilia at the tips of the hair cells which lies within the organ of Corti on the basilar membrane, and this leads to action potentials (spikes) within the nerve fibers of the auditory nerve. Because each inner hair cell is attached to a specific place on the basilar membrane, each neuron in the auditory nerve carries information about the vibration of the basilar membrane at a single place in the cochlea. This means each neuron in the auditory nerve is sensitive to each characteristic frequency.

The auditory nerve carries the information about incoming sound from the cochlea to the cochlea nucleus. Cells of the cochlear nucleus project to higher nuclei through ventral and dorsal streams (Pickles 2008). Cells that project via a ventral stream primarily project to the superior olivery complex (SOC). The SOC is divided into three primary nuclei: the medial superior olive (MSO), lateral superior olive (LSO), and the medial nucleus of the trapezoid body (MNTB). The timing and intensities of the stimuli at the two ears are compared in the SOC, and the information is used for sound localization. Cells that project via a dorsal stream project mainly to the contralateral lateral lemniscus nuclei (LMN) and inferior colliculus (IC). They are involved in the complex analysis of a sound.

The ventral stream, mainly involved in sound localization, runs primarily to the SOC. The stream has two divisions. In the first division, the intensities of the stimuli at the two ears are compared in the LSO. In the second division, the timing of the stimuli in the two ears is compared in the MSO. The dorsal stream, mainly involved in sound identification, runs primarily to the IC of the opposite side, some fibers synapsing in the LMN on the way.

The IC is the main receiving station for the ascending pathways from lower stages of the brain stem. It forms the primary site of convergence of the sound identification and sound localization streams. It is suggested that this is a critical stage in transformation from responses dominated by the simple acoustic characteristics, to those which integrate acoustic properties in a way that begins to define an auditory object (Pickles 2008).

The medial geniculate body (MGB) is the specific thalamic auditory relay of the auditory system, receiving afferent from the IC, and projecting to the cerebral cortex. It also has heavy reciprocal connections back from the cortex, indicating that the cortex and MGB are grouped together as a functional unit.

The auditory cortex consists of core areas, surrounded by belt and parabelt areas. Previous studies suggest that the core area is necessary for the response to basic features of a sound, while the belt and parabelt areas are necessary for the response to complex features. It is suggested that the auditory cortex is necessary for the representation of auditory objects, that is, the assembly of information about all features of a sound (Pickles 2008).

To evaluate sound qualities, it is necessary to consider our auditory functioning, that is, how incoming sound is processed from the peripheral to central auditory system.

2.2 Neural Evidences of the Autocorrelation Model in the Auditory Pathways

A model for evaluations of environmental noise has been proposed based on human auditory system (Ando 2001). The model of the auditory-brain system includes the autocorrelation function (ACF) mechanism, which might exist in the auditory nerve, and the interaural cross-correlation function (IACF) mechanism, which might exist in the IC. Temporal and spatial sensations may be processed in the left and right hemisphere according to the temporal factors extracted from the ACF and the spatial factors extracted from the IACF, respectively, which is discussed in Chaps. 3 and 4. The overall subjective responses, for example, subjective preference and annoyance, may be processed in both hemispheres in relation to the temporal and spatial factors (Ando 2002).

The internal ACF may provide a representation for judging pitch salience. In temporal models of pitch perception, it is assumed that the pitch is extracted with autocorrelation (Licklider 1951; Bilsen and Ritsma 1969; Wightman 1973; Yost and Hill 1979; Meddis and Hewitt 1991; Patterson et al. 1995). Regarding the pitch salience or strength of pitch, psychophysical research has revealed that the strength of the pitch corresponds well to the peak amplitude of the ACF of the auditory signal, which represents the degree of temporal regularity of the sound (Wightman 1973; Yost et al. 1996; Yost 1996; Ando et al. 1999). One type of sound that allows for systematic manipulation of pitch salience is iterated rippled noise (IRN). IRN is produced by adding a delayed copy of a noise to the original noise and then

repeating this delay-and-add process (Bilsen 1966; Yost 1996). A normalized ACF of IRN reveals a peak at the reciprocal of the delay, whose magnitude grows with increasing number of iterations reflecting the increasing periodicity.

Physiologically, recordings of responses to IRN stimuli from auditory nerve fibers (Fay et al. 1983; ten Kate and van Bekkum 1988) and cochlear nucleus neurons (Bilsen et al. 1975; Shofner 1991, 1999; Winter et al. 2001) show that the pitch of IRN is represented in the firing patterns of action potentials locked to either the temporal fine structure or envelope periodicity. That is, there is temporal regularity in the fine structure of the neural firing patterns, and it produces peaks in the autocorrelogram. These data suggest that the pitch of IRN stimuli is based on ACF mechanism. Indeed, the pooled interspike interval distributions of auditory nerve discharge patterns in response to complex sounds resemble the ACF of the stimulus waveform, and the magnitude of the ACF peak corresponds well with pitch salience (Cariani and Delgutte 1996a, b).

Electroencephalography (EEG), the measurement of electric potential differences on the scalp, is a widely applied method for investigating the functions of the human brain. Magnetoencephalography (MEG) is closely related to EEG. In both methods, the measured signals are generated by the same synchronized neuronal activity in the brain. The time resolution of EEG and MEG is in the millisecond range. Thus with EEG and MEG, it is possible to follow the rapid changes in cortical activity that reflect ongoing signal processing in the brain; the electrical events of single neurons typically last from one to several tens of milliseconds (Hämäläinen et al. 1993).

Several types of stimulus-evoked brain stem neural activity may be recorded using the EEG. Best known and most extensively studied EEG is the auditor brain stem response (ABR). Another type of brain stem neural activity is the frequency-following response (FFR). Unlike the ABR, the FFR reflects sustained neural activity (integrated over a population of neural elements) that is phase-locked to the individual cycles of the stimulus waveform and/or the envelope of periodic stimuli (Krishnan 2007).

The ABR and auditory evoked magnetic field (AEF) were recorded and analyzed to identify such ACF mechanism in human brain. FFRs were recorded from seven listeners in response to IRNs, which varied only in their degree of pitch salience (Krishnan et al. 2010). The FFR reflects sustained phase-locked activity in a population of neural elements within the rostral brain stem (e.g., Krishnan 2007; Chandrasekaran and Kraus 2010).

To create IRN stimuli with a dynamic fundamental frequency (F0) contour whose pitch varies as a function of time, a time-varying delay-and-add algorithm to a filtered Gaussian noise (10–3000 Hz) was applied (Denham 2005). The pitch increases in a curvilinear fashion from about 100 to 135 Hz over the 250-ms stimulus duration. By using a different number of iterations (n) in the IRN generating circuit, the F0 contour's pitch salience was varied. FFRs were recorded from each listener in response to monaural stimulation of the right ear at 80 dB sound pressure level (SPL) through a magnetically shielded insert earphone (Etymotic, ER-3A) (Krishnan et al. 2010). Neural responses were recorded differentially

between a noninverting (+) electrode placed on the midline of the forehead at the hairline (Fz) and inverting electrodes (−) placed on the left (A1) and right (A2) mastoid, and the 7th cervical vertebra (C7). Another electrode placed on the mid-forehead (Fpz) served as the common ground. FFRs were recorded simultaneously from the three different electrode configurations and subsequently averaged for each stimulus condition to yield a response with a higher signal-to-noise ratio.

To analyze the robustness of encoding to stimuli differing in pitch salience, the neural pitch strength of each response waveform was quantified. From each FFR, the normalized ACF calculated over the entire duration of the response was computed in order to determine the dominant periodicities contained within the response. The height of the first peak in the ACF from time-delay zero was taken as the magnitude of neural pitch strength (Krishnan et al. 2005). In all cases, this peak fell at a delay of 10 ms, the fundamental pitch period of the input stimulus (Fig. 2.5). The growth in FFR pitch strength (derived from peak magnitude of the FFR ACF) with increasing iteration steps suggests an increase in the degree of neural phase-locking to the pitch relevant periodicity resulting from increased temporal regularity in the stimulus.

In addition, behavioral frequency difference limens (F0 DLs) were measured from each listener to obtain a perceptual estimate related to pitch salience. F0 DLs decreased with increasing stimulus periodicity revealing better pitch change detection for more salient stimuli. The strong correlation observed between the neural and behavioral measures supports the view that ACF-related pitch encoding at a subcortical and sensory level of processing plays an important role in shaping pitch perception.

MEG has been used to investigate how features of sound stimuli related to pitch are represented in the human auditory cortex. Focused on a spatial representation of pure tone in the auditory system according to their frequency, tonotopic organization of the human auditory cortex has been investigated (e.g., Elberling et al. 1982; Romani et al. 1982; Pantev et al. 1988, 1995). Focusing on the temporal structure of the sound, the periodicity pitch-related cortical response has been investigated (Pantev et al. 1989; Langner et al. 1997; Cansino et al. 2003; Fujioka et al. 2003; Seither-Preisler et al. 2003).

AEFs in relation to bandwidth variations of band-pass noise have been examined (Soeta et al. 2005a, 2006). The results indicate that the peak amplitude of N1m, which is found above the left and right temporal lobes around 100 ms after the stimulus onset, decreases with increasing bandwidth of the band-pass noise. The peak amplitude of the ACF increases with decreasing bandwidth of the auditory stimuli (Merthayasa et al. 1994; Sato et al. 2002; Soeta et al. 2004a). The pitch strength of band-pass noises was found to increase with decreasing bandwidth (Fastl and Stoll 1979). Therefore, it was assumed that sounds that have larger peak amplitude of the ACF or stronger pitch could lead to more cortical activity, which would cause an increase in the strength of the N1m response.

To evaluate responses related to pitch salience, which is characterized by the peak amplitude of the ACF of the sound, in auditory cortex, the AEFs elicited by

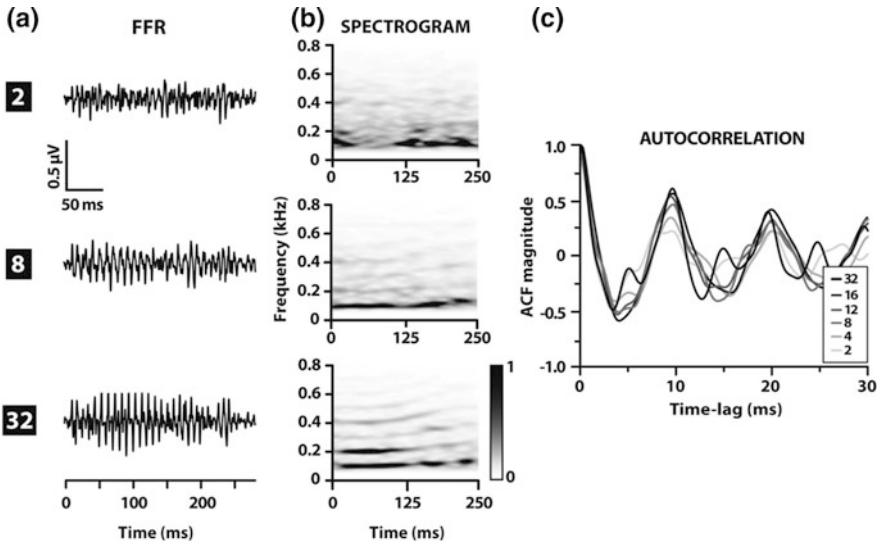


Fig. 2.5 FFR waveforms (**a**), spectrograms (**b**), and ACFs (**c**) as a function of iteration steps (n) computed from grand averaged brain stem responses. Spectrograms were computed by sliding a 50-ms analysis window by a 1 ms increment and computing the FFT in each time bin. Normalized magnitudes are indicated by the *gray-scale* gradient; *darker shades* indicate stronger encoding of pitch relevant harmonics (i.e., more pronounced phase-locked activity). Minimal periodicity is observed in FFRs at low iteration steps ($n = 2$; *top row*). By $n = 8$ iterations, FFR phase-locked activity captures periodicity related to the fundamental frequency (F0) and its harmonics (*middle row*). Robust encoding is even more pronounced at $n = 32$ iterations when the stimulus is maximally salient (*bottom row*). Temporal waveforms and ACFs derived from the FFRs reveal increasing periodicity in the neural response with increasing iteration steps, thus indicating more robust brain stem activity for salient pitch (Krishnan et al. 2010)

IRN with different iteration numbers were recorded (Soeta et al. 2005b). It was anticipated that the N1m amplitude would increase with an increase in the number of iterations of the IRN.

Ten normal hearing listeners (22–31 years, all right-handed) took part in the experiment. The IRN was produced by a delay-and-add algorithm applied to band-pass noise that was filtered using fourth-order Butterworth filters between 400 and 2200 Hz. The number of iterations of the delay-and-add process was set at 0, 1, 4, and 16. The delay was fixed at 1 ms, corresponding to a pitch of 1000 Hz. The stimulus duration used the experiment was 0.5 s, including rise and fall ramps of 10 ms. The auditory stimuli were delivered to the listeners through plastic tubes and inserted earpieces at a comfortable listening level adjusted separately for each listener. Figure 2.6 shows the temporal waveforms and the power spectra of some of the stimuli measured with an ear simulator that includes a microphone, a pre-amplifier, and an adaptor connected to the earpiece. Figure 2.7 shows the ACF of some of the stimuli measured with the ear simulator. The τ_1 of IRN corresponds to the delay. The ϕ_1 increases as the number of iterations increases.

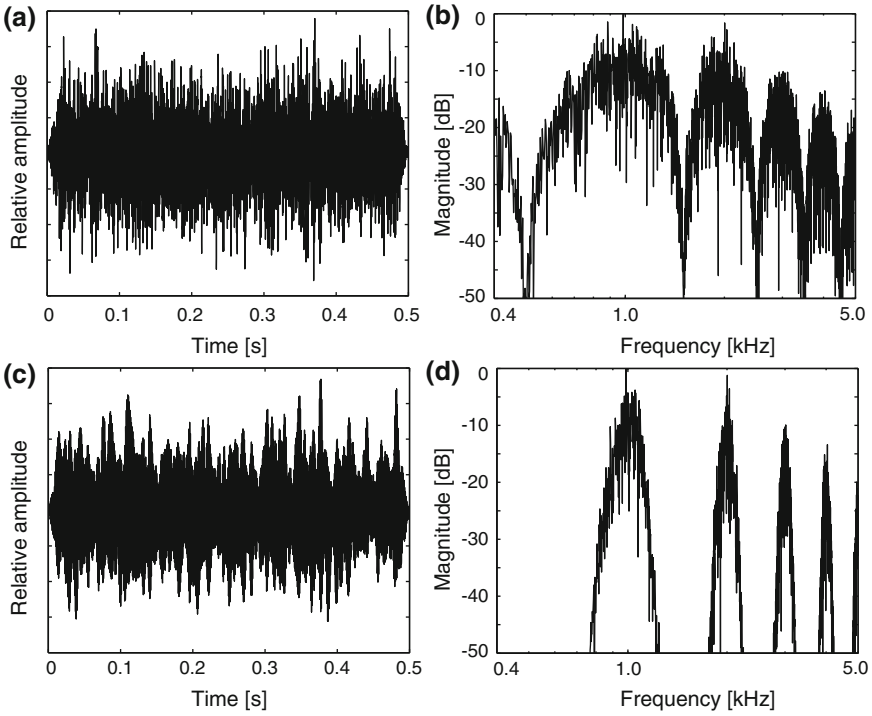


Fig. 2.6 Temporal waveforms and power spectra of the IRN with the number of the iterations (**a**, **b**) 1 and (**c**, **d**) 16 (Soeta et al. 2005b)

The AEFs were recorded using a 122 channel whole-head DC superconducting quantum interference device (DC-SQUID) magnetometer (Neuromag-122^{TM}; Neuromag Ltd., Helsinki, Finland) in a magnetically shielded room (Hämäläinen et al. 1993). The IRNs were presented in randomized order with a constant inter-stimulus interval of 1.5 s. Listeners were instructed to watch a self-selected silent movie and to ignore the stimuli. The magnetic data were sampled at 0.4 kHz after being band-pass filtered between 0.03 and 100 Hz and then averaged approximately 50 times. The averaged responses were digitally filtered between 1.0 and 30.0 Hz. The analysis time was 0.7 s from 0.2 s prior to the stimulus onset, with an average prestimulus period of 0.2 s serving as the baseline. The Neuromag-122^{TM} has two pick-up coils in each position that measure two tangential derivatives, $\partial Bz/\partial x$ and $\partial Bz/\partial y$, of the field component Bz. With such a coil configuration, the largest signal occurs just above a dipolar source where the gradient is steepest (Knuutila et al. 1993). To evaluate the amplitude and latency of the N1m peak, the root-mean-squares (RMS) of $\partial Bz/\partial x$ and $\partial Bz/\partial y$ were determined as the amplitude of the responses at each recording position. The N1m peak amplitude and latency were defined as the RMS peak and latency in the latency range from 70 to 130 ms over

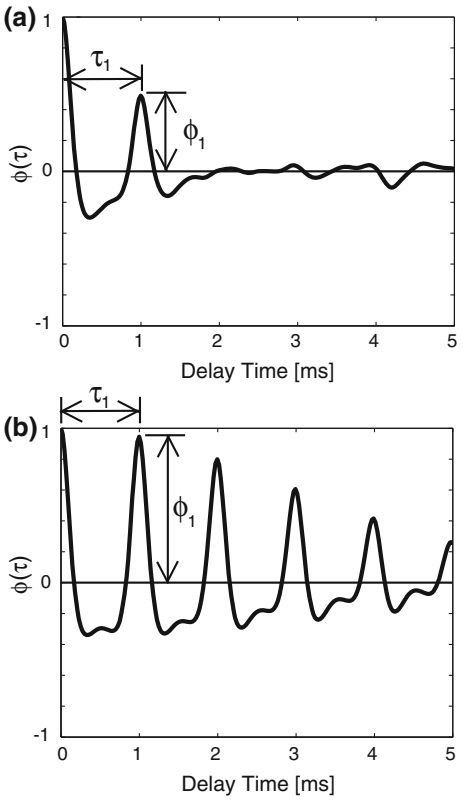


Fig. 2.7 ACFs of the IRN with the number of the iterations **a** 1 and **b** 16 (Soeta et al. 2005b)

the right and left hemispheres. In each listener, we employed the N1m peak latency and amplitude with a channel that showed the maximum amplitude placed at each hemisphere.

To estimate the location and strength of the underlying neural activity of the N1m wave, a single equivalent current dipole (ECD) was assumed as the source of the magnetic field of the N1m wave in a head-based coordinate system. The ECDs that best described the measured magnetic field at the N1m peak latencies were found by least squares fitting in a spherical volume conductor (Kaukoranta et al. 1986). A one-dipole model was used separately for the left and right hemispheres, with a subset of channels over each hemisphere. The origin of this coordinate system was set at the midpoint of the medio-lateral axis (y-axis) which joined the center points of the entrance to the acoustic meatuses of the left and right ears. The posterior–anterior axis (x-axis) was oriented from the origin to the nasion, while the inferior–superior axis (z-axis) was perpendicular to the x–y plane. All ECDs with goodness-of-fit values exceeding 80 % were used in further analyses.

Clear N1m responses were observed in both the right and left temporal regions in all listeners (Fig. 2.8). The N1m latencies were not systematically affected by the number of iterations of the IRN. Figure 2.9 depicts the mean N1m amplitude

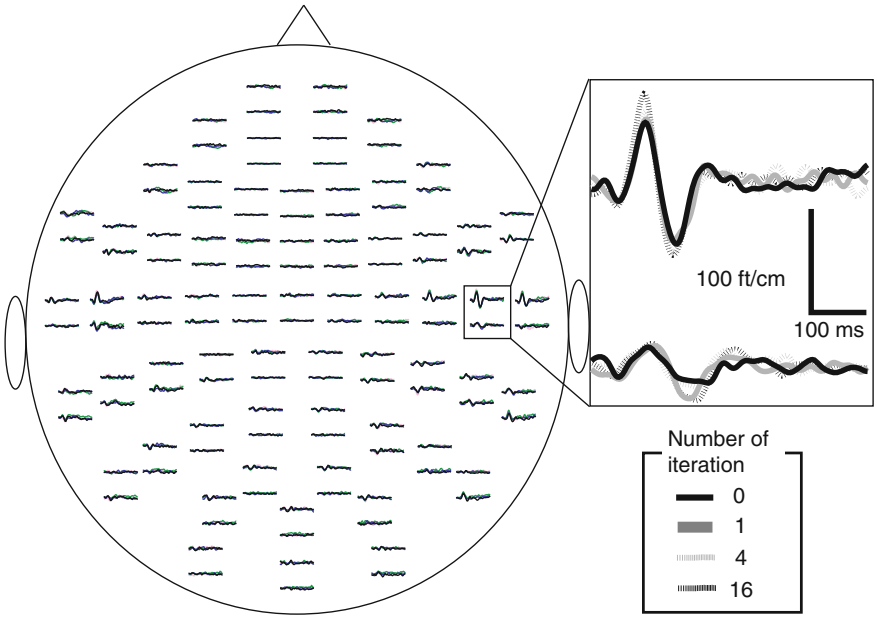


Fig. 2.8 Typical waveforms of auditory evoked magnetic fields from 122 channels in a listener (Soeta et al. 2005b)

(across ten listeners) as a function of the number of iterations. A greater number of iterations of the IRN produced a larger N1m amplitude, that is larger ϕ_1 of the stimulus produced a larger N1m response. A two-way analysis of variance (ANOVA) (number of iterations × hemisphere) revealed a significant main effect of the number of iterations on the N1m peak amplitude. This result is consistent with the previous study using band-pass noise with variation of bandwidth (Soeta et al. 2005a, 2006) and IRN with a delay of 2, 4, or 16 ms (Krumbholz et al. 2003; Soeta and Nakagawa 2008a). The amplitude of the AEF component in response to periodic stimuli was compared with simulated peripheral activity patterns of the auditory nerve (Seither-Preisler et al. 2003). The results showed that the amplitude of the N1m is correlated with the pitch strength estimated on the basis of auditory nerve activity. This result is consistent with the present results.

For the dipole strength, similar results to those for the N1m peak amplitude were obtained. A greater number of iterations of IRN produced a larger N1m ECD moment. A two-way ANOVA revealed a significant main effect of the number of iterations of IRN on the ECD moments. The ECD locations did not show any systematic variation across the listeners as a function of the number of iterations of IRN.

Figure 2.10 shows the relationship between ϕ_1 of the stimulus and ECD moment of the N1m response. To compare the previous result, the results of band-pass noise with center frequency of 1000 Hz were also included in Fig. 2.10 (Soeta et al. 2006). Psychophysical studies have indicated that first peak of the ACF, ϕ_1, could

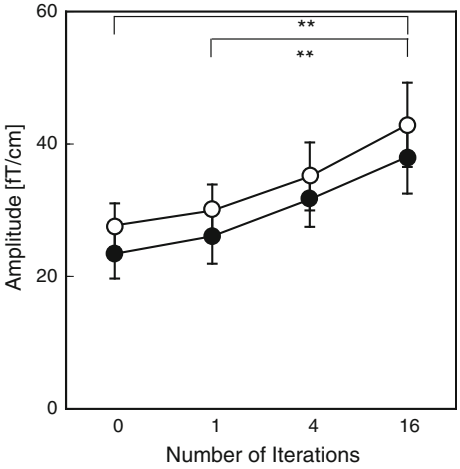


Fig. 2.9 Mean amplitude of the N1m (±SEMs) as a function of the number of iterations from the right (*filled circle*) and left (*open circle*) hemispheres. The *asterisks* indicate statistical significance (*$P < 0.05$, **$P < 0.01$; Post hoc Bonferroni test) (Soeta et al. 2005b)

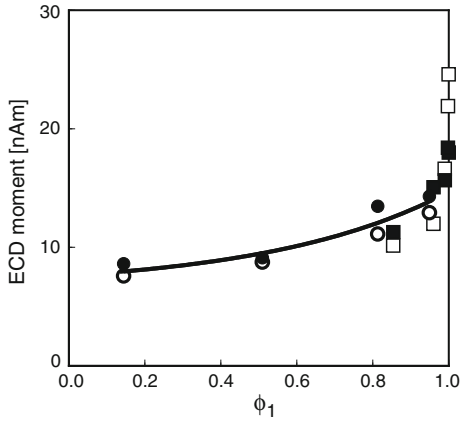


Fig. 2.10 Relationship between peak amplitude of the ACF, ϕ_1, and ECD moment of the N1m from the right (*filled circle*) and left (*open circle*) hemispheres. The curve is of the form $6.6 + 100.9*\phi_1$ (Soeta et al. 2005b). The previous ECD moment of the N1m in response to band-pass noise with center frequency of 1000 Hz from the right (*filled square*) and left (*open square*) hemispheres were also included for comparison (Soeta et al. 2006)

account for the pitch strength of the stimulus (Wightman 1973; Yost et al. 1996; Ando et al. 1999) and pitch strength of IRN was an exponential function of ϕ_1, namely of the form $a + 10b\phi_1$, where "a" and "b" are constants (Yost 1996). Note that, the ECD moment derived in the current study could also be described in the form $a + 10b\phi_1$, where $a = 6.6$ and b = 0.9, as shown in Fig. 2.10.