Duration sensation in relation to the autocorrelation function of Sound stimuli

Kazi, Saifuddin

DURATION SENSATION IN RELATION TO THE
AUTOCORRELATION FUNCTION
OF SOUND STIMULI

(音源信号の自己相関関数に関連する時間感覚)

THE DISSERTATION IS SUBMITTED

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE SCHOOL OF SCIENCE AND TECHNOLOGY

KOBE UNIVERSITY

(神戸大学)

By

KAZI SAIFUDDIN

July 2001

GENERAL PREFACE

The dissertation is submitted, in partial fulfillment of the requirements for the **Doctor of Philosophy** degree, at the Graduate School of Science and Technology, **Kobe University**.

The basic theme of the dissertation is to investigate the empirical relationship between the temporal phenomena of the material existence and its human sensitivity. In the whole study temporal phenomena were determined by the factors extracted from the autocorrelation function of the sound stimuli. On the other hand, sensitivity of the temporal phenomena is defined by the differential subjective judgments of the auditory duration.

I would like to express my deep appreciation to **Professor Yoichi Ando** for his considerable guidance and constant suggestions. I take the responsibility of any unknown errors in the pages occurred.

Kobe   **Kazi Saifuddin**

July 2001

ACKNOLEDGMENTS

ABSTRACT

A series of studies was conducted as the partial fulfillment of the requirements for the Doctor of Philosophy degree in the discipline of psychoacoustics. To examine the subjective reality of temporal duration in the primary sensation of sound stimuli psychophysical test methods were used. The properties of sound stimuli were determined by the factors extracted from its autocorrelation function (ACF). Using the "model of the auditory-brain system" with the ACF, and the "theory of primary sensation" the below theory is concluded.

"Duration sensation of the sound stimuli is one of the primary sensations of the human-brain neuron which is responsible for the factors τ_1, τ_e, ϕ_1, ϕ_m and $\Phi(0)$ extracted from the ACF and physical duration."

Which can be expressed by the equation:

$$DS = f[\tau_1, \tau_e, \phi_1, \phi_m, \Phi(0), PD].$$

Here, DS: duration sensation; PD: physical duration; and f : function. τ_1: delay time of the first peak; τ_e: effective duration of the envelope at the ten-percentile delay; ϕ_1: amplitude of the first peak; ϕ_m: constant value of all amplitudes; and $\Phi(0)$: energy at the origin in the ACF. Smaller symbols signify minor effects on DS.

CONTENTS

CHAPTER 3. DS IN RELATION TO THE FACTOR $\phi_m(\phi_m = \phi_1; m = 2, 3, ...)$ OF ACF FOR THE MIXED SITIMULI OF PURE-TONE AND WHITE-NOISE

CHAPTER 4. DS IN RELATION TO THE FACTORS τ_1, ϕ_1 AND τ_e OF THE AC F OF BANDPASS-NOISE STIMULI

CHAPTER 5. DS IN RELATION TO THE FACTORS τ_1, ϕ_1, τ_e AND $\Phi(0)$ OF THE ACF OF BANDPASS-NOISE STIMULI UNDER REVERBERATION CONDITIONS

CHAPTER 1

INTRODUCTION

1.1. Preface

The temporal dynamism is become one of the most premising research phenomena since the modern civilization began. Because of the belief that no change can take place without specification of time. Since the remarkable development of the philosophy and science, time is defined differently with many unsolved controversies. Isaac Newton in his *Philosophiae Naturalis Principia Mathematica* (1687) quoted that "absolute, true, and mathematical time, of itself and from its own nature, flows equably without relation to anything external". He also explained the underlying idea in psychophysics that "subjective reality is a direct reflection of objective reality". Concluding the modern era of the philosophy and science, time can be recognized with the various aspects as psychological time, biological time, solar time, mechanical time, atomic time, etc. The twentieth century has been emphasized on time, explaining it as the fourth dimension of space in the renowned "Theory of Relativity" of Albert Einstein (Martin et al. 1995).

At the beginning of the twenty-first century psychological time may be considered as the emergence of the new era of science and philosophy. The reality and fantasy whatever, and however, the time as absolutely continuos or conditional with the space, the then reflection of brain function should be studied more. In the cognitive and

neuroscience, the evaluation of the sensation of temporal duration within some milliseconds of mechanical time can bring a significant result if it is multiplied with the broader time billions of year. Considering the human life as a continuous process from individual to individual and from generation to generation through the biological evolution and the sensation as a psychological evolution over the sequence of physical time, the present study can be significant.

1.2. Previous studies

Fraisse described the history of remarkable researches on the perception of time in his book "The Psychology of Time" in 1963 and made a comment that "time is a subject with a crowded research history over the 20th century". The human experience of the time duration of a sound signal has been studied from various point of view. Gyuau (1902) found that estimates of time depend on many variables, such as the number, type, and intensity of sound stimuli presented and the subject's expectations, levels of attention, and interest. Woodraw (1951) described his data in terms of "Vierordt's law": that shorter time intervals are overestimated and longer intervals are underestimated. He also described an indifferent range in which intervals are neither overestimated nor underestimated. Frankenhaeuser (1959) presented clicks as acoustical stimuli, at various intervals from 13 to 72 s and sequenced at various speeds, and found that estimates of duration increased with the speed at which the clicks were sequenced. Fraisse (1961) found that an interval during which signals are presented at 5/s is estimated to be longer than an interval in which they are presented at 10/s. An early explanation of duration

discrimination was proposed by Creelman (1962) that a central counter is postulated that sums the number of random neural firings (counts) that occur during the duration to be judged. Michon (1965) described the variation of subjective duration with the type and amount of activity in which the subjects are engaged. In 1966 Austin et al. used auditory clicks at frequencies between 0 and 10/s and found that the magnitude of the frequency effect was a monotonically decreasing function of duration. Vroon (1970) found that the apparent durations increase with the rate of information transmission and that when the subject has to behalf actively, i.e., by translating the stimuli into binary choices, the experienced duration decreases with the number of processed bits. Sharon (1971) used two silent durations (T and $T + \Delta T$) in a two-alternative forced-choice procedure and found that discrimination depended on the parameters of the marker. Huggins (1972) explained his experimental results that subjects are much more sensitive to changes in vowel duration than to changes in consonant duration. He also added that the perception of timing in natural speech is based on events at the syllabic level rather than at the segmental level. Hicks at al. (1977) stated that the time appears to pass more rapidly if one is engaged in a higher level of behavioral activity. Ando (1977) found that the time passes more quickly when it is filled with noise than when it is silent and explained this to an internal clock inhibited by noise. Zakay et al. (1983) found that subjective estimates of time were a decreasing function of task difficulty, and that estimates of the durations of "empty" intervals were longer than those of "filled" intervals. Subjects produced the longest estimates of duration when the external tempo was fastest and produced shorter estimates when the external tempo were slower. Kato et al. (1992,

1993) found that perceptual sensitivity to the change of duration is dependent on the intensity using speech segments as stimuli. Paster and Artieda (1996) stated that the internal clock can be imagined to be an oscillating circuit, with a single pacemaker neuron or a group of pacemaker neurons, that generates rhythmical electromagnetic activity the brain uses in the temporal analysis of information.

Ando et al. (including author) conducted one questionnaire study with the author in 1999 to examine the temporal experiences of longer duration (see Ando et al. 2000a). Results indicated that both subjective time durations of one day and three years are well described by the normal distribution if the time scale is logarithmic rather than linear. A remarkable finding obtained here is that, in reference to the mean subjective time duration of the junior high school (three years), the mean subjective time durations for each of the three-year periods before the elementary school were significantly longer than unity (about 1.2), and the subjective duration of the three-year period of the senior high school is significantly shorter (about 0.8). When subjects transferred between elementary school at 6-11 years of age, the mean subjective time duration of the periods below 11 years of age was much longer, more that 1.3.

1.3. Model of auditory-brain system

A model of auditory-brain system was proposed by Ando (1998) as shown in Figure 1.1. It may explain well the subjective attributes and the auditory evoked potentials even though the continues brain waves corresponding to the changing phenomena of acoustic signals. In this figure, sound source is the $p(t)$ and r_o indicates it's three-dimensional

position. The sitting location of r means the center position of the head of listener, $h_{l,r}(r\,|\,r_o,t)$, being the impulse responses between r_o and the entrances of left and right ear-canal. The impulse responses of the external ear canal and the chain of three bones (hammar, stirup, anvil) indicate $e_{l,r}(t)$ and $c_{l,r}(t)$, respectively (The structure of the human ear is shown as a cross-section in the Figures 1.2). The movements of the basiler membrane are expressed by $V_{l,r}(x,\omega)$ in which, x being the position alone the membrane. The hair cells produce action potentials those are transmitted through the cochlear nuclei, the superior olivary complex up to the auditory cortex of both hemispheres. The superior olivary complex includes the function of medial superior olive, the lateral superior olive and the trapezoid body.

The input power density spectrum of the cochlea $I(x')$ can be roughly mapped, according to the tuning of a single fiber (Katsuki et al., 1958 and Kiang, 1965) at a certain nerve position x'. This fact may be partially supported by auditory brainstem response (ABR), waves (I -IV) which reflect the sound pressure levels as a function of the horizontal angle of incidence to a listener (Ando et al. 1991). The sound pressure level, which corresponds to the denominator of Equation (1.1) with the ACFs for the two ears at the origin of time ($\sigma = 0$) which, appears in the latency at the inferior colliculus, may be processed in the right hemisphere. These neural activities, in turn, include necessary information to attain the ACF at the higher level of the lateral lemniscus as indicated here by $\Phi_{ll}(\sigma)$ and $\Phi_{rr}(\sigma)$. It is known that the left hemisphere is mainly associated with time-sequential identifications, and the right is concerned with spatial identifications (Kimura, 1973; Sperry, 1974; Ando, 1987). But
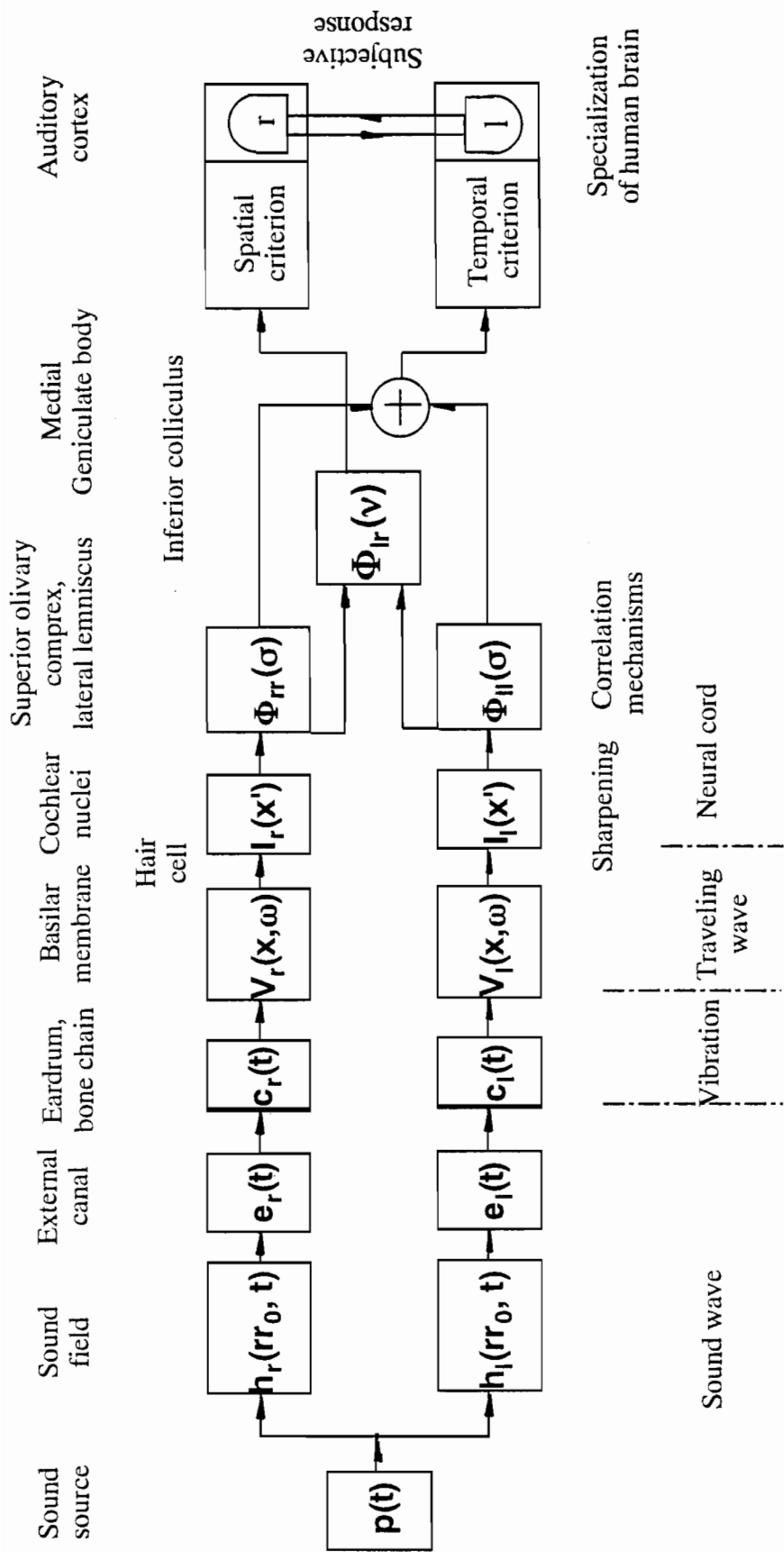
Figure 1.1. An auditory-brain model for subjective responses.

the hemispheric dominance is a relative phenomenon depending on what factor is changed in the presented pair.
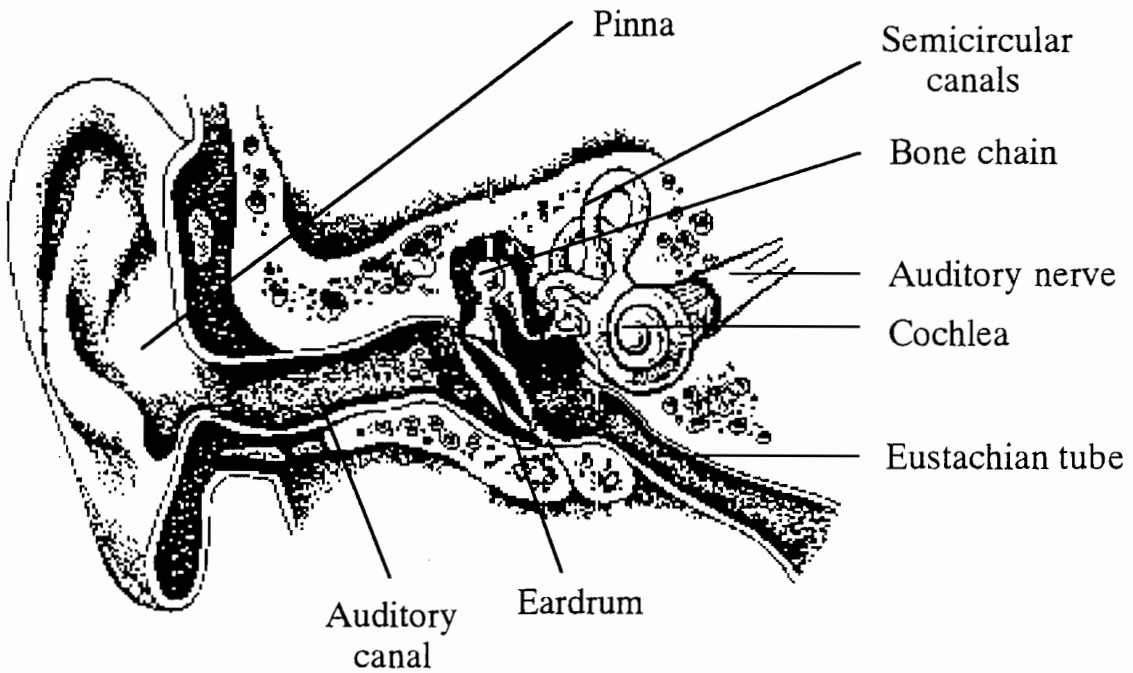
Figure 1.2. Diagram of the entire ear. The outer, middle, and inner ear are shown, along with adjacent structures.

The normalized interaural cross-correlation function is defined by

$$\phi_{lr}(\tau) = \frac{\Phi_{lr}(\tau)}{\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}},$$ (1.1)

where $\Phi_{ll}(0)$ and $\Phi_{rr}(0)$ are ACFs at $\tau = 0$ for the left and right ear, respectively, or the sound energies at both ears. Here, the denominator is

$$\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}$$

According to this model subjective attributes are described by all the acoustic factors

associated with both cerebral hemispheres. The echo disturbance, the threshold of perception of a reflection and coloration (temporal information) are considered, to be dominantly processed in the left hemisphere related to the envelope function of the ACF of sound stimuli.

1.4. Factors of autocorrelation function

Among the most useful signal processors concerned with the auditory system autocorrelation function (ACF) is one of them. Such ACF can be defined by

$$\Phi_p(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} p'(t)p'(t+\tau)dt, \tag{1.2}$$

where $p'(t) = p(t) * s(t)$, $s(t)$ being the ear sensitivity. To use practically convenient, $s(t)$ may be selected as the impulse response of an A-weighted network. So the ACF can be obtained from the power density spectrum, which is defined by

$$P_d(\omega) = P(\omega)P^*(\omega), \tag{1.3}$$

where $P(\omega)$ is the Fourier Transform of $p(t)$, which is the particular signal. That is why the ACF can be expressed by

$$\Phi_p(\tau) = \int_{-\infty}^{+\infty} P_d(\omega)e^{j\omega t}d\omega, \tag{1.4}$$

$$P_d(\omega) = \int_{-\infty}^{+\infty} \Phi_d(\tau)e^{-j\omega t}d\tau, \tag{1.5}$$

Thus it can be recognized that the ACF and the power density spectrum contain the same mathematical information. The normalized autocorrelation function (NACF) is defined by

$$\phi_p(\tau) = \frac{\Phi_p(\tau)}{\Phi_p(0)},$$

(1.6)

The signal duration in the ACF corresponding to the psychological present, as suggested by Fraisse (1982), is 2T (integration interval) = 0.5- 5.0 s. Such of psychological present defined here as the short-time duration of stimuli needed for subjective responses. Analyzing the ACF, five significant factors can be found and are defined as below.
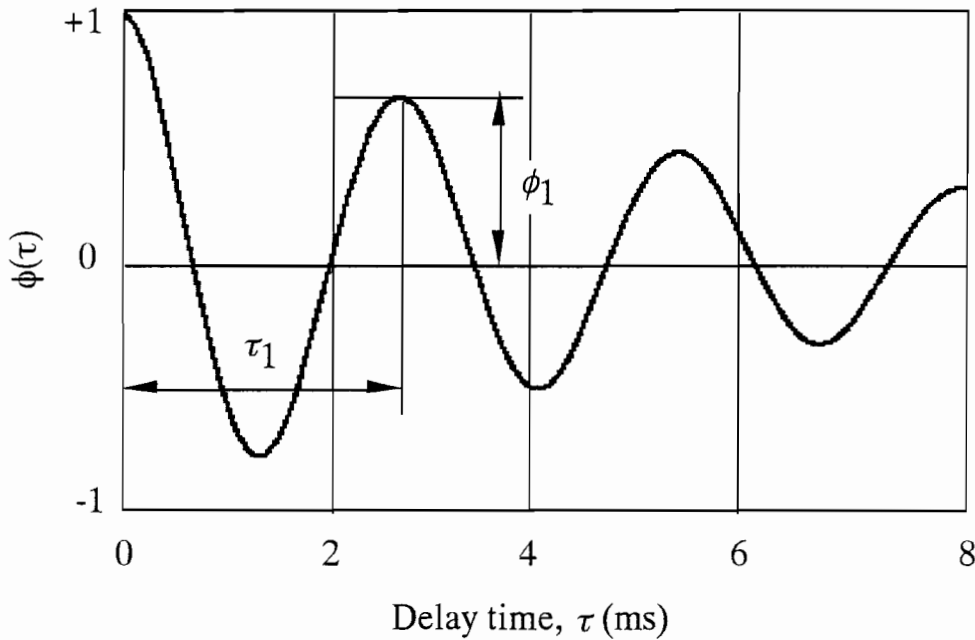
Figure 1.3. The τ_1 and ϕ_1 are defined here by the delay time and amplitude, respectively of the first peak of the ACF.

1. The time delay of the first peak, τ_1, in the normalized ACF is defined by Equation (1.6) and is shown in Figure 1.3.

2. The effective duration of the envelope of the normalized ACF, τ_e, is defined by the ten-percentile delay or the -10 dB, representing a kind of repetitive feature or reverberation contained within the source signal itself (Figure 1.4).
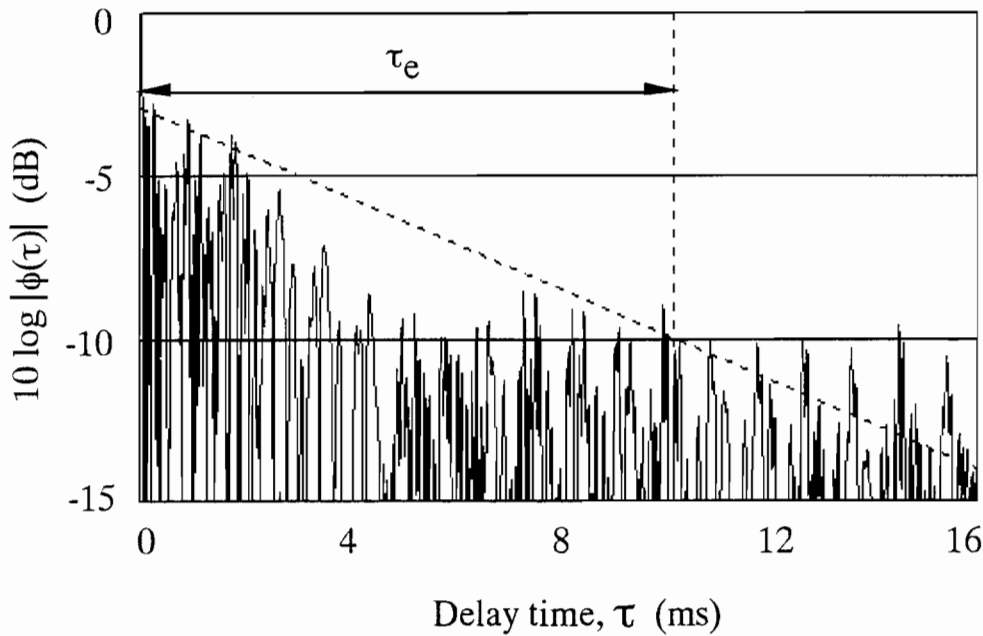
Figure 1.4. An example of determining τ_e. The τ_e is defined by the delay time at which the envelope of the ACF becomes -10 dB.

3. The amplitude of the first peak, ϕ_1, in the normalized ACF is defined by Equation (1.6) and is shown in the Figure 1.3.

4. The value of amplitude ϕ, in which, ϕ_1, ϕ_2, ϕ_3, . . . , ϕ_n, are found constant in the ACF and specified as ϕ_m and is shown in Figure 1.5.

5. The energy represented at the beginning of the time delay $\Phi_p(\tau)$, here the $\tau = 0$. Such energy is indicated here by $\Phi(0)$.
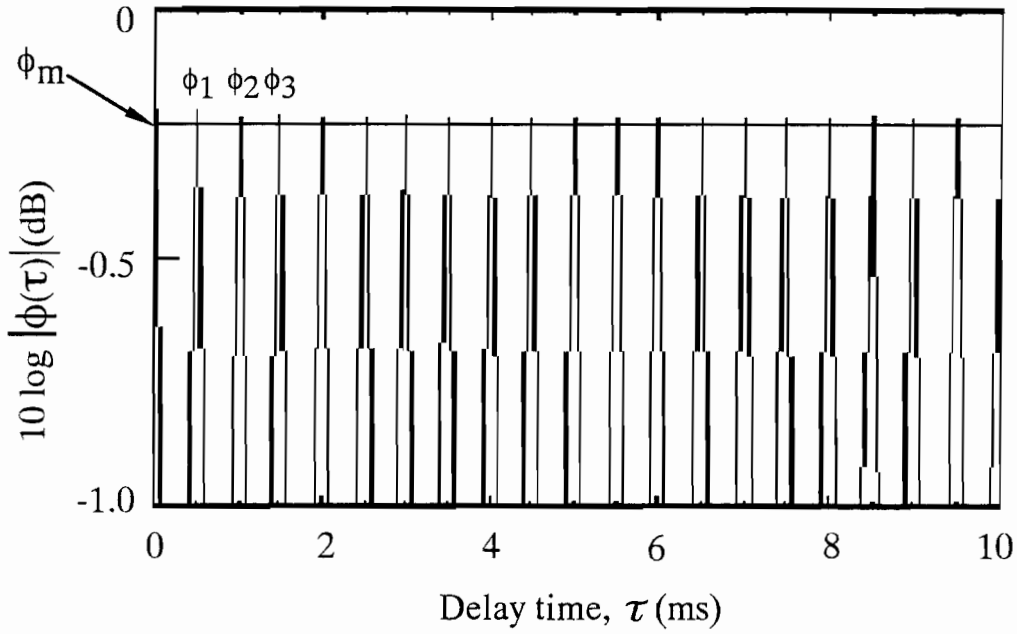
Figure 1.5. The ϕ_m is defined by the constant value of the peaks (ϕ_1, ϕ_2, ϕ_3, $\cdots$, ϕ_n) of the ACF. The ACF is measured for the stimulus of pure-tone (1000 Hz) mixed with white-noise.

1.5. Primary sensation and autocorrelation function

The model of auditory-brain system (Ando, 1998) can be supposed as working well to describe fundamental subjective attributes for sound field. It also describes the autocorrelation function (ACF) and interaural cross-correlation function (IACF), according to the Figure 1.1. Another theory corresponding to the auditory-brain system was proposed by Ando (2001a and b) to explain the fundamentals of sensation of the auditory stimulus. That is the "theory of primary sensation" of sound stimulus. If $c_i(i = 1, 2, ..., I)$ are to be the physical factors representing cues influencing any primary sensations $S_i(i = 1, 2, ..., J < I)$, a sensation s_j may be then express by

$$s_j = f(c_1, c_2, ..., c_i), \quad j = 1, 2, ..., J. \tag{1.7}$$

On the other hand, when the physical factors are orthogonal, then s_j may be expressed by a linear combination, then

$$s_j = f(c_1) + f(c_2) + ... + f(c_i), \quad j = 1, 2, ..., J. \tag{1.8}$$

Ando described in the same study (2001a) whether or not the sensations are independent of each other. That can be easily demonstrate by the case as

$$s_j = f_j(c_1) + f_j(c_2), \quad s_k = f_k(c_1) + f_k(c_2), \tag{1.9}$$

the correlation coefficient in between s_j and s_k can be given by

$$r_{jk} = \overline{s_j s_k} = f_j(c_1)f_k(c_1) + \overline{f_j(c_2)f_k(c_2)} + \overline{f_j(c_1)f_k(c_2)} + \overline{f_j(c_2)f_k(c_1)}. \tag{1.10}$$

In this equation the first and second terms of the right-hand side are not always zero. It means that such primary sensations are subject to being not mutually independent. The described primary sensation can be found differently as expressed by the subjective attributes. Four primary sensations of sound stimuli are briefly explained below.

a) **Loudness:** Considering the first primary one, sensation of loudness (s_L) may be defined by (Ando 1998; Merthayasa et al. 1994; and Sato et al. 2001)

$$s_L = f_L[\Phi(0), \tau_1, \phi_1, \tau_e], \tag{1.11}$$

where different values of τ_1 determine the pitch or missing fundamental.

b) **Pitch:** The second one is the sensation of pitch (s_p) which, can be given by (Sato et al. 2001, Wightman F. L. 1973, Inoue et al. 2001)

$$s_p = f_p(\tau_1, \phi_1), \tag{1.12}$$

c) **Timbre:** The sensation of timbre (s_T) is the third primary sensation which is the complicated one, can be differentiate by the psychological attributes within the same

loudness, pitch and duration (Plomp 1970). It may be expressed by

$$s_T = f_T[\Phi(0), \tau_e, (\tau_1, \phi_1), ..., (\tau_n, \phi_n)]. \qquad (1.13)$$

where (τ_1, ϕ_1) are the significant orthogonal factors in (τ_n, ϕ_n), n = 1, 2, . . . , equation (1.12) can be expressed as

$$s_T = f_T[\Phi(0), \tau_e, \tau_1, \phi_1]. \qquad (1.14)$$

d) **Duration:** The sensation of duration (s_D) is the fourth one of the four primary sensations. The present study is going to find out the empirical relations between s_D (In the whole study s_D is used as DS) and the factors extracted from the ACF.

1.6. Purposes of the study

The previous studies showed that the human perception of temporal duration is not at all absolute as like as the physical time duration. The perceptual time is obviously conditional with both subjective and objective factors. It is also known that the space is three-dimensional and any change of the three dimensional structure is determined by the time as a fourth one. To understand the changes of the material structure temporal phenomena must be considered. If the material-phenomena of three dimensions are reflected by the brain temporal phenomena should be included. It is obvious that the human brain is functioning at the very developed hierarchy level of biological evolution, which however, includes it's primitive characters. The subjective judgment of the temporal duration of sound stimuli is described here as duration sensation (DS) on the basis of the 'theory of primary sensation'. In order to this to understand the primitive functions of temporal sensation of the neuron of human brain the study was conducted.

To achieve such of goal the below objectives were specified.

1. To examine the relationships between the duration sensation (DS) and the factor τ_1, being the time delay of the first peak of the autocorrelation function (ACF).

2. To examine the relationships between the DS and the factor τ_e, being the effective duration of the ACF in which the envelope become -10dB.

3. To examine the relationships between the DS and the factor ϕ_1, being the amplitude of the first peak of the ACF.

4. To examine the relationships between the DS and the factor ϕ_m, being the constant value of amplitudes of all the peaks in the ACF. The factor ϕ_m also includes ϕ_1.

5. To examine the relationships between the DS and the factor $\Phi(0)$, being the energy at the beginning of the ACF.

6. To examine whether or not the DS is related to that of the PD (physical duration).

1.7. Methods of subjective test

Allan (1979) classified the different methods of measuring time perception into two major categories: duration scaling and duration discrimination. In duration scaling tasks, subjects are asked about the perceived duration of a set of easily discriminating temporal intervals, while in duration discrimination tasks, subjects are requested to distinguish among a set of confusable intervals. However, duration-scaling tasks include temporal production, synchronization, magnitude estimation, temporal reproduction,

and category rating. While duration discrimination tasks include comparison, single stimulus, and time bisection.
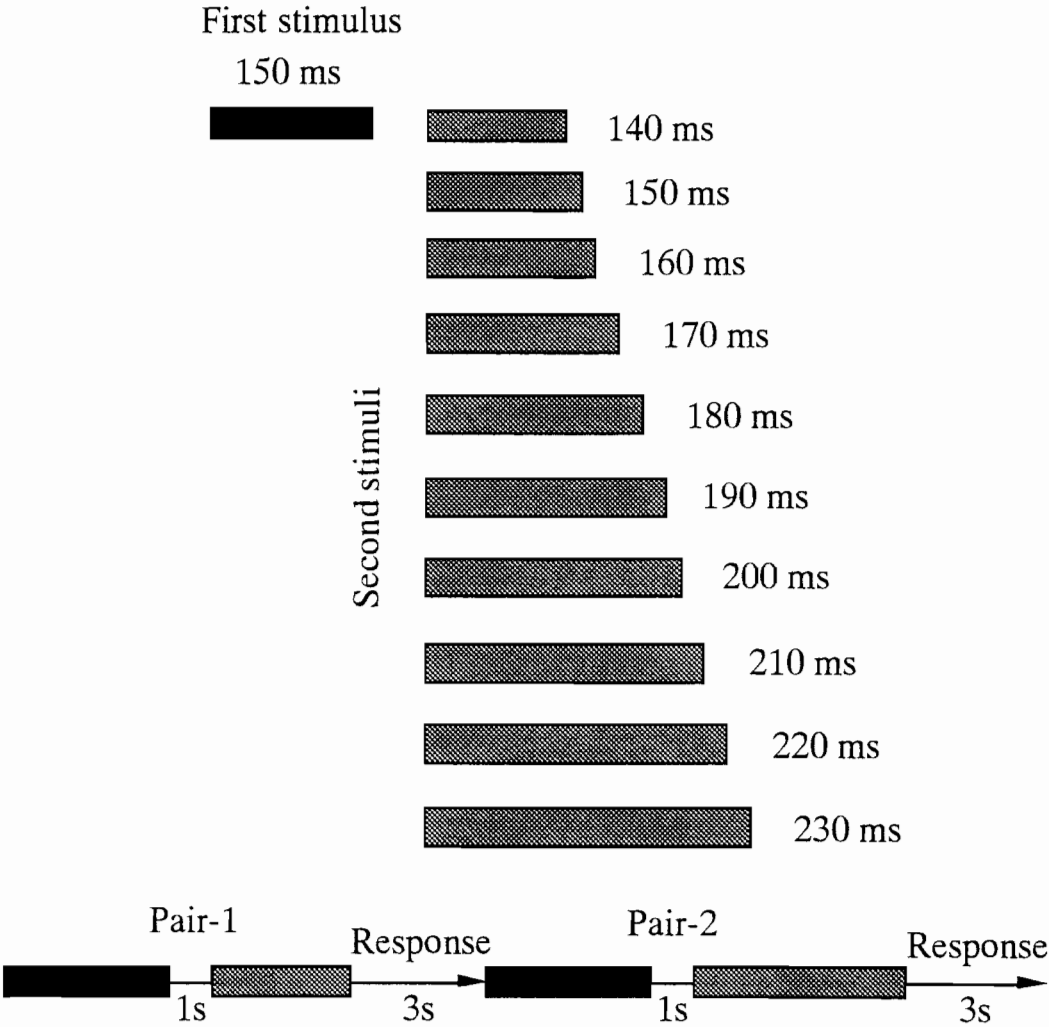
Figure 1.6. Diagram of paired-comparison test indicates the black-part as pure-tone stimulus used first in the pair with the constant duration and gray-part as white-noise stimuli with different durations. Intre-pair and inter-pair gaps indicate 1 s and 3s, respectively.

In this study paired-comparison method was used to simplify the subjective

judgments. The paired-comparison method usually needs a number of judgments for a single pair (Ando 1998; 141-143). Regarding this an excellent discussion is given in Robinson et al. (1973) and Hall (1983). In this case, two stimuli were presented in each trial, and the subject is asked whether the duration of the second is longer than, shorter than, or equal to the first with the respect of same parameter in each pair. Different pairs contain different parameters. Usually first stimulus remains constant and the second one is varied trial to trial, randomly (see Figure 1.6).
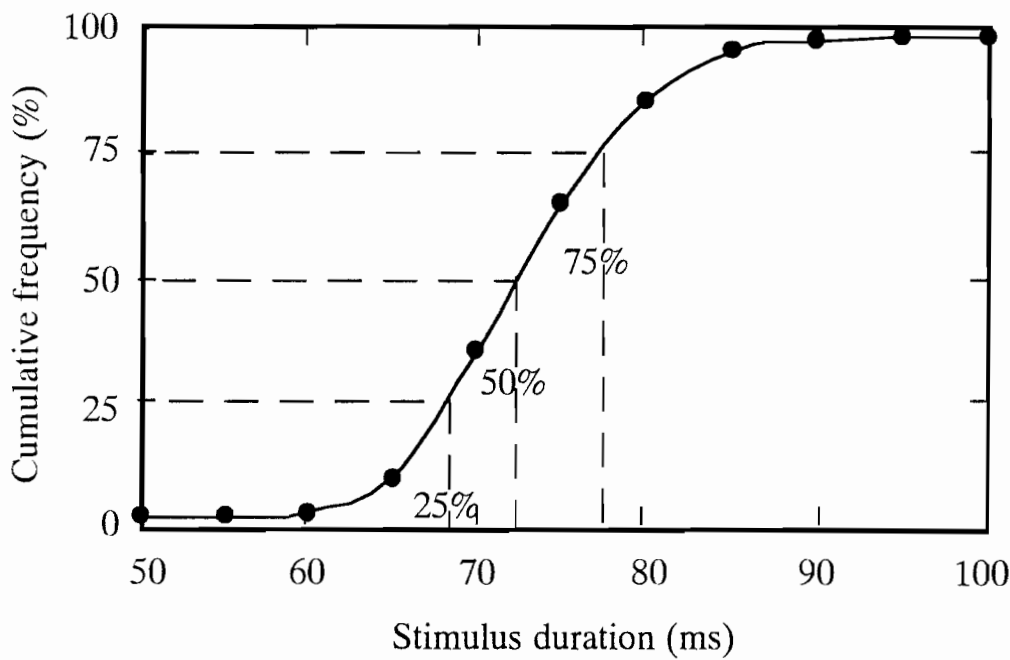
Figure 1.7. An example is shown to explain the differential sensitivity of the data obtained by the paired-comparison test. The 25%, 50%, and 75% are the different levels of the function.

The subject has to response if the second stimulus seems longer than the first. The cumulative frequency of "longer" judgments is simply obtained by counting the

percentage of the responses. The response is counted as "correct judgment" when it is obtained by pushing particular switch for the expected judgment. An example of the correct judgments for the one parameter experiment is shown in the Figure 1.7 in which, three levels of duration discrimination are obtained. The 50% line of the duration judgments is defined here as the duration sensation (DS).

The paired-comparison test also used here to obtain scale values in the experiment 2, Chapter 3 using "A method of calculating individual subjective responses by paired-comparison tests" by Ando and Singh (Ando and Singh; 1996). The method is based on the Thurstone's model (Thurstone, 1927), and the scale values are calculated within the linear range of the plot of probability. Different responses are obtained on the longer judgments for the two stimuli in the pair. The individual data are confirmed by using a "test of goodness of fit". A discussion of procedure for calculating scale values and for using the test of goodness of fit of duration judgment is shown in the APPENDIX - A.

CHAPTER 2

DURATION SENSATION (DS) IN RELATION TO THE FACTOR τ_1 OF AUTOCORRELATION FUNCTION (ACF) OF PURE-TONE AND COMPLEX-TONE STIMULI

2.1. Preface

It is well known that the sound associated with sinusoidal waves is called a pure tone. Mathematically the pure tone can be described by the equation

$$x(t) = A\sin(2\pi t / T + \phi) \tag{2.1}$$

where A is amplitude, T is the period in seconds, and ϕ is the phase in radians. The units of x, whatever they might be, are the same as the units of the amplitude A.

Such of sound wave is frequently used as the standard because of its convenient frequency distribution. To understand the neuron responses of auditory stimulus pure tone can be used as standard form. Galambos and Davis (1943) stated on the responses of single auditory nerve fibers to acoustic stimulation at various frequencies. They explained the auditory response areas of single neurons as a function of frequency of the sound stimulus. A narrowly tuned cell responds to a very limited range of frequencies, whereas a broadly tuned cell responds to a much wider frequency range. Since an unstimulated neuron maintains an ongoing spontaneous discharge rate even in the absence of any apparent stimulation, is threshold may be determined by varying the

stimulus level until the lowest intensity is reached at which the neuron responds above

its spontaneous rate. An alternative approach is to present the stimulus at a fixed

intensity, and to measure the number of spike potentials with which the unit responds at

different stimulus frequencies. The former method measures the neuron's sensitivity,

and the latter is firing rate, as a function of frequency. The frequency with the lowest

threshold (or the greatest firing rate) is the characteristic frequency (CF). Such CFs of

individual auditory nerve fibers are related to distance alone the length of the cochlea

(Liberman, 1982). Kiang (1965) found that the discharges of auditory nerve fibers are

time locked to tonal stimuli up to 4000-5000 Hz. This relation was demonstrated by the

presence of the post-stimulus time (PST) histogram of single peaks corresponding to

individual cycles of the stimulus. It is also consistent with other evidence that auditory

nerve fibers respond to the particular phase of the stimulus within the frequency range

(Hind et al. 1967; Rose et al. 1967).

In order to this, many of the significant studies show that the frequency of the sound

stimuli plays an important role on the so-called neural coding of the brain function. In

this study frequency of the sound stimulus was described as the factor of the

autocorrelation function (ACF). That is the time delay of the first peak of the ACF,

which can be expressed by τ_1.

A new model of the auditory-brain systems concerning with the autocorrelation

function and interaural cross-correlation mechanisms of sound stimuli is described in

1998 by Ando. Using this model he proposed another 'theory of primary sensation'

(2001a). In this theory four primary sensations are recognized as loudness, pitch, timbre,

and duration. Thus DS in the auditory system is well concerned with the factors extracted from the autocorrelation function (ACF) as a signal process which, is supposed to be equivalent to the power density spectrum of the sound signal at the entrances of the ears.

The pitch is the psychological correlate of frequency, such that high frequency tones are heard as being "high" in pitch and low frequencies are associated with "low" pitches. According to the American National Standards (ANSI), standard of 1994, "pitch is that attribute of auditory sensation in terms of which sound may be ordered on a scale extending from low to high". Pitch depends mainly on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus. The perception of pitch is concerned with the factor τ_1, as the time delay of the first peak in the ACF, which can be determined by the frequency (Ando et al. 1998, 2000b, 2001a; Flanagan et al. 1958). It was found that the larger values of τ_1 or lower pitch in the ACF of the pure-tone and bandpass-noise stimuli (with and without reverberation effects) produce longer DS. If the frequency of pure-tone stimulus is concerned with the DS, missing fundamental of complex-tone stimulus also can be examined here (Seebeck 1841; Schouten et al. 1962). This study used pure-tone and complex-tone stimuli consisting of harmonic components without the fundamental frequency (missing-fundamental) which is called residue pitch (Schouten et al.1940). The conducted study includes three experiments (1, 2 and 3)

In the experiment 1, the DS was examined in relation to τ_1 of the ACF of pure-tone stimuli with the frequency of 125, 250, 500, 1000, 2000 and 4000 Hz. In the

experiment 2, the *DS* was examined in relation to τ_1 of the ACF of complex-tone stimuli with three fundamental frequencies (250, 500 and 1000 Hz). In both experiments (1 and 2), stimuli were compared with white-noise duration. In another experiment 3, *DS* was examined on an comparison directly between the complex-tone and pure-tone stimuli. The complex tone by the f_6 and f_7 components (fundamental frequency 500 Hz) was compared with the two pure-tone stimuli of 500 and 3000 Hz.

2.2. Methods

2.2.1. Stimuli preparing

In the experiment 1, six pure-tone stimuli with different frequencies of 125, 250, 500, 1000, 2000 and 4000 Hz were selected to use in this study. The duration of the stimuli was 150 ms constant. Ten white-noise stimuli with a 10 ms step size durations in the range from 140 to 230 ms were selected. The rise and fall times of all the stimuli were 1 ms constant. The rise and fall times were defined as the time taken to reach a level −3 dB different from the steady level.

In the experiment 2, three complex-tone stimuli (Figure 2.1) followed by the three fundamental frequencies (f_0) of 250, 500, and 1000 Hz were selected to use in the three subjective-test sessions. All of the three stimuli were consisted of the two components. The first and second harmonics of the three fundamental frequencies are shown in the Table 2.1. During construction of complex-tone wave in the digital computer the starting phases of the two harmonic components was adjusted at zero. Ten white-noise stimuli were selected with the ten durations of 140 to 230 ms maintaining 10 ms step

size as like as experiment 1.

Table 2.1. Three complex-tone stimuli consisted of the two pure-tone components which are harmonics of the three fundamental frequencies (f_0).

Fundamental frequency f_0 (Hz)	First component (Hz)	Second component (Hz)
250	3000	3250
500	3000	3500
1000	3000	4000

In the experiment 3, one complex-tone stimulus with the fundamental frequency of 500 Hz (Table 2.1) and two pure tone components (f_6: 3000 and f_7: 3500 Hz) were selected. All of the stimuli included constant rise and fall times of 1 ms (same to experiment 1).

2.2.2. ACF measurement

The autocorrelation function (ACF) of all the pure-tone stimuli for the experiment 1 were analyzed (2T = 2 s) and found the values of the τ_1 (delay time of the first peak) are 8.0, 4.0, 2.0, 1.0, 0.5 and 0.25 ms, respectively, with the frequency of 125, 250, 500, 1000, 2000 and 4000 Hz. The ACFs were measured from the recorded stimuli in the sound proof chamber. The effective duration τ_e of the ACF was found almost constant (parallel to the steady level).
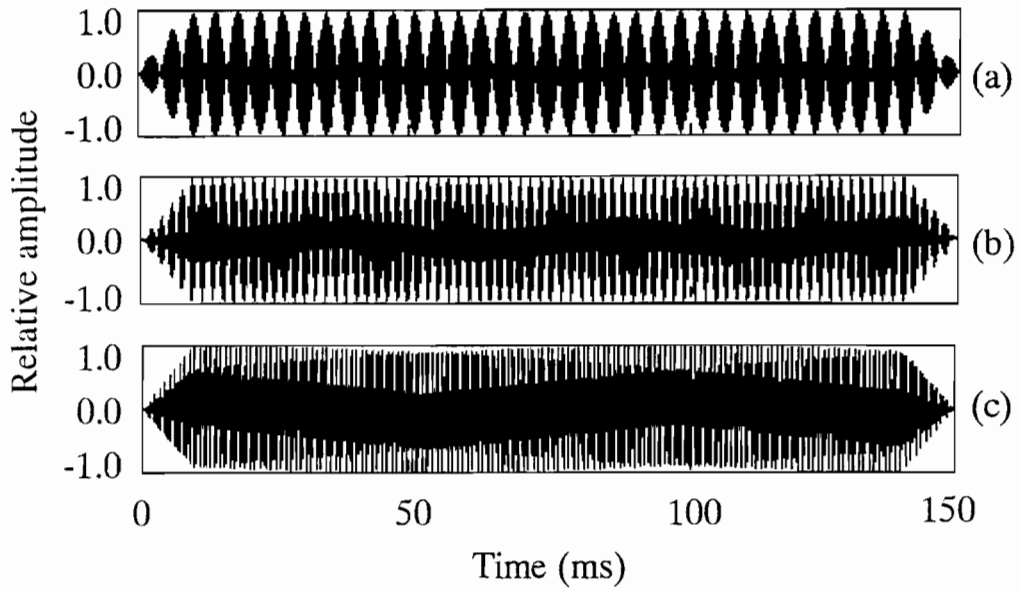
Figure 2.1. Real wave-form of three complex-tone stimuli (a), (b), and (c) indicate the different fundamental frequencies (f_0) respectively, 250, 500 and 1000 Hz. Second components of the complex-tone stimuli are 3250, 3500, and 4000 Hz, respectively, and the first component is a 3000 Hz constant. Total duration 150 ms for each stimulus includes 1 ms rise and fall times.

The ACF of the three complex-tone stimuli was measured for the experiments 2 and 3 in the sound proof chamber of which the integration interval ($2T$) was 2 s (Figure 2.2). The amplitudes of the peaks (ϕ_1, ϕ_2, ϕ_3, . . .,ϕ_n) in the ACF of all stimuli were found almost parallel. The time delay of the first peak (τ_1) of the ACF of complex-tone stimuli were found same as to the missing fundamentals. The ACF of pure-tone stimulus (500 and 3000 Hz) was also measured.

2.2.3. Subjects

The subjects of the experiment 1 were five students in the laboratory (KF, MI, NS, OK and YS). In the experiments 2 and 3, ten subjects were participated, respectively, JA,

TK, KF, RS, NA, TH, KS, NK, YO, NA; and MK, DG, SK, MN, KA, NK, DB, NA, MA, SS. All of them have good hearing and were between 22 and 25 years old.
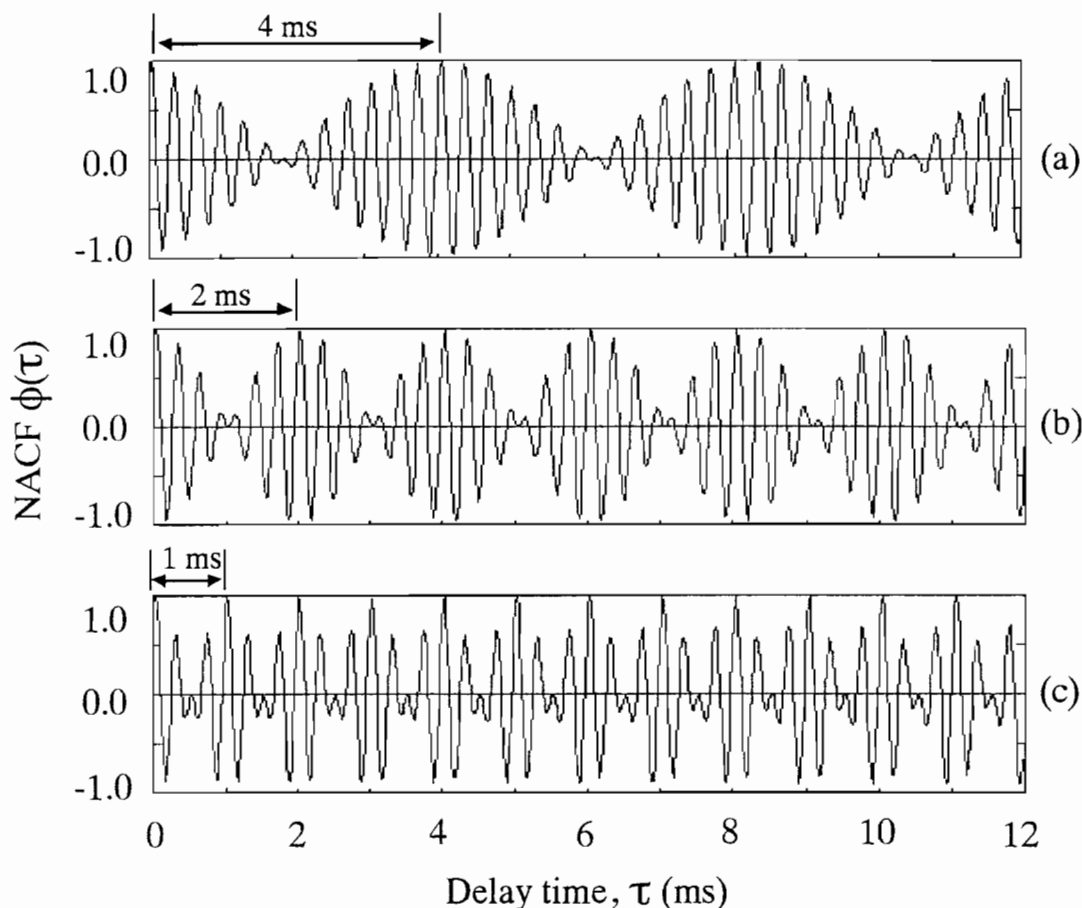
Figure 2.2. Normalized autocorrelation function (NACF) of three complex-tone stimuli (a), (b), and (c) indicate the different fundamental frequencies (f_0) respectively, 250, 500, and 1000 Hz. The second components are 3250, 3500, and 4000 Hz, respectively and the first component is a 3000 Hz constant.

2.2.4. Stimuli presentation

In the experiment 1, using paired-comparison test six pure-tone stimuli were presented in the six different sessions using different frequencies (as measured by τ_1).

Ten sets of white-noise stimuli with a 10 ms step size in the range from 140 to 230 ms were used at second in the pair for each session. In the paired-comparison test, the second stimuli were presented in random order. Each pair was presented 20 times. The intra-pair and inter-pair (response time) gaps, respectively, were 1 and 3 s. The sound pressure level (SPL) of all stimuli was kept at a constant 80 dB(A) [A: being an weighted network].

In the experiment 2, using paired-comparison test, three complex-tone stimuli were presented first in the pair followed by three sessions with constant 150 ms duration. Otherwise procedure were same to that of the experiment 1.

In the experiment 3, paired-comparison tests were used to obtain the scale values (SV) for the comparison of duration sensation (DS) between pure-tone and complex-tone stimuli (for procedure see APPENDIX – A). One complex-tone and two pure-tone stimuli were used in this study. Three durations (140, 150 and 160) for each stimulus out of three were used in the pair maintaining matrix combinations (N(N-1)/2, N = 9). The parameter was the pitch in relation to the τ_1 of the pure-tone and complex-tone stimuli. Otherwise methods of pair-presentations were almost same to that of the experiments 1 and 2.

For all of the three experiments, source stimuli were presented from a single loudspeaker at a horizontal distance of 74 (± 1) cm with the center of the head of the seating listener in the soundproof chamber. The sound pressure levels (SPL) of all the stimuli used in the study were kept constant 80 dB(A). The chamber was dark at the time of taking experiment.
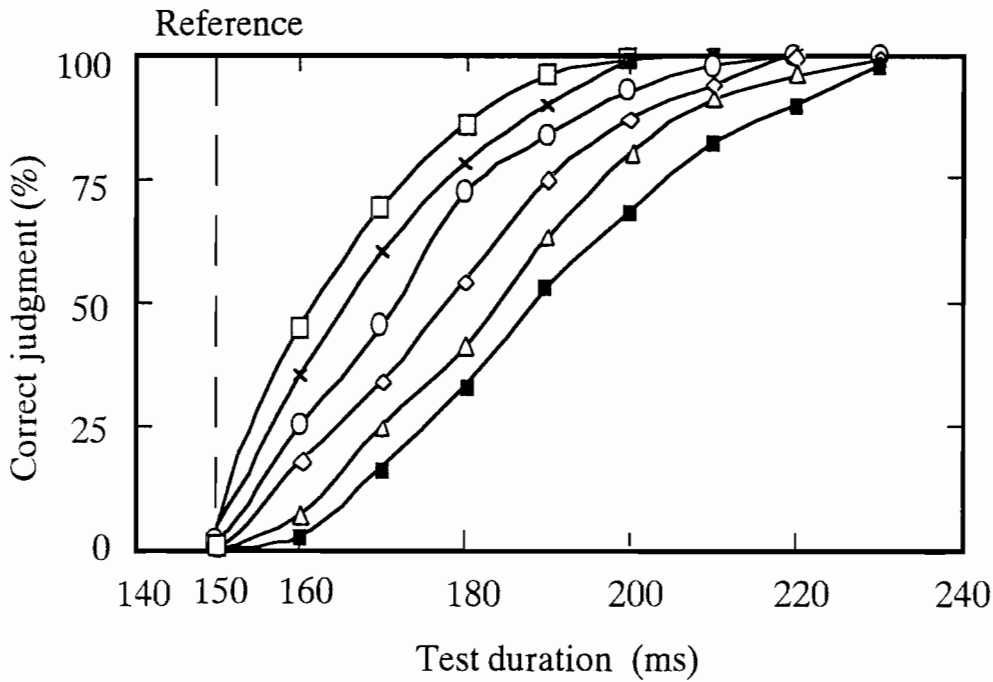
Figure 2.3. Cumulative frequencies of correct judgments indicate the 50% line of duration for six stimuli as ■: 125; △: 250; ◇: 500; ○: 1000; ✕: 2000; and □: 4000 Hz. Varticle-dashed line indicates the physical duration of the pure-tone stimulus (150 ms).

2.2.5. Subjective test

The subjects were seated in the soundproof chamber on a chair in front of a loudspeaker. They were asked to judge whether the duration of the first or second of a pair of stimuli (the test stimuli) was longer than or shorter than that of the other (Stanley, 1998). Throughout the experiments 1 and 2, subjects were told to push the button only when the second stimulus (white noise) seemed longer than the standard (first stimulus). Each subject went through six sessions with six different test stimuli in the experiment 1 and three sessions with three stimuli in the experiment 2. In the experiment 3, subjects had to push the button using two options whether the first or the second stimulus seems longer than that of the another in the pair. Nine stimuli were used through nine different

sessions. Between the two sessions the rest time was 5 minutes.

2.3. Results

2.3.1. Experiment 1

Cumulative frequencies of the correct judgments are obtained as duration for the five subjects by the effect of six different frequencies (controlled by τ_1) of a pure tone (see Figure 2.3). The pure-tone stimulus with a constant duration (150 ms) was presented at first for comparison with a second white-noise stimulus in the pair. Correct judgments of duration obtained for the six stimuli cross the 50% correct lines are found at 162.0, 165.5, 172.0, 177.8, 184.0 and 188.0 ms, respectively. The percentages of the judged-longer duration than the standard duration (150 ms) are found respectively, 8.0 %, 10.3 %, 14.7 %, 18.5 %, 22.7 %, and 25.3 %. Differences among the durations judgement of the 6 stimuli were found to be significantly ($p < 0.01$) different. Individual differences of the judged values of duration are shown in the Figure 2.4 and the standard deviation of those scores was 11.9 ms. The average of the individual data is shown in Figure 2.5. In Figure 2.5, the relation between DS and τ_1 is well described by

$$DS \approx \alpha(\log \tau_1) + \beta, \qquad (2.2)$$

here, $\alpha \approx 15$ and $\beta \approx 10$.

Results of the analysis of variance (ANOVA) of individual data for the judged durations are listed in Table 2.2.
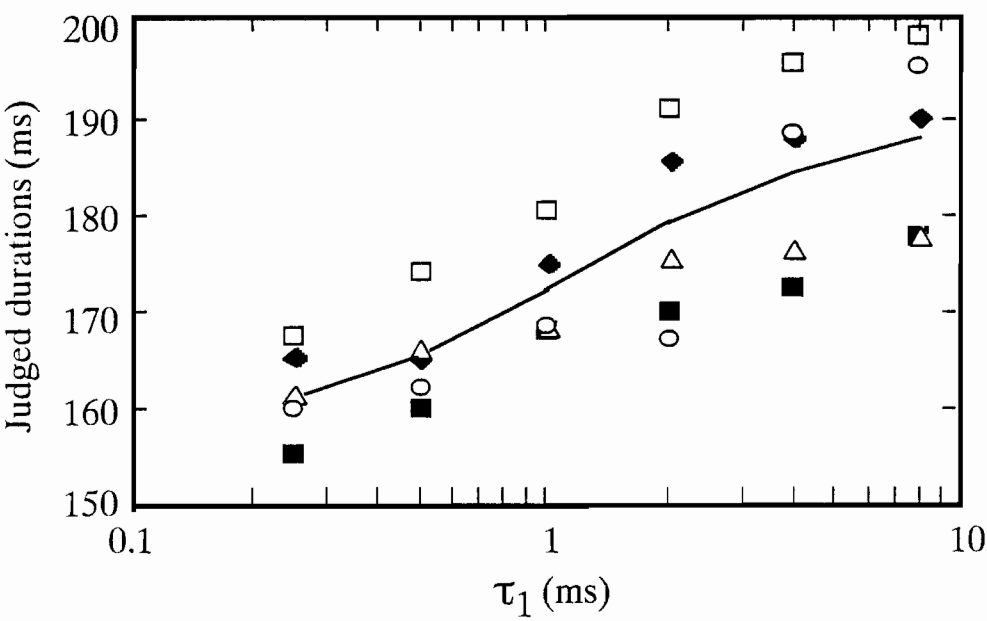
Figure 2.4. Individual differences of the judged duration for pure-tone stimuli with the six different frequencies (inversely express by the τ_1 in the ACF). The symbols indicate the different scores of the five subjects ($\square$: KF; $\blacklozenge$: MI; $\bigcirc$: NS; $\blacksquare$: OK; $\triangle$: YS; and —: average).

Table 2.2. F-values of the analysis of variance for the judged-durations among the six stimuli of pure tone by different frequencies were compared with white-noise stimulus.

Frequency (Hz)	125	250	500	1000	2000	4000
Frequency (Hz)						
125	—	4.02	10.24*	15.96**	22.48**	26.00**
250		—	6.22	11.94*	18.46**	21.98**
500			—	5.72	12.24*	15.76**
1000				—	6.52	10.04*
2000					—	3.52
4000						—

* and ** indicate 5% and 1% significant level, respectively.
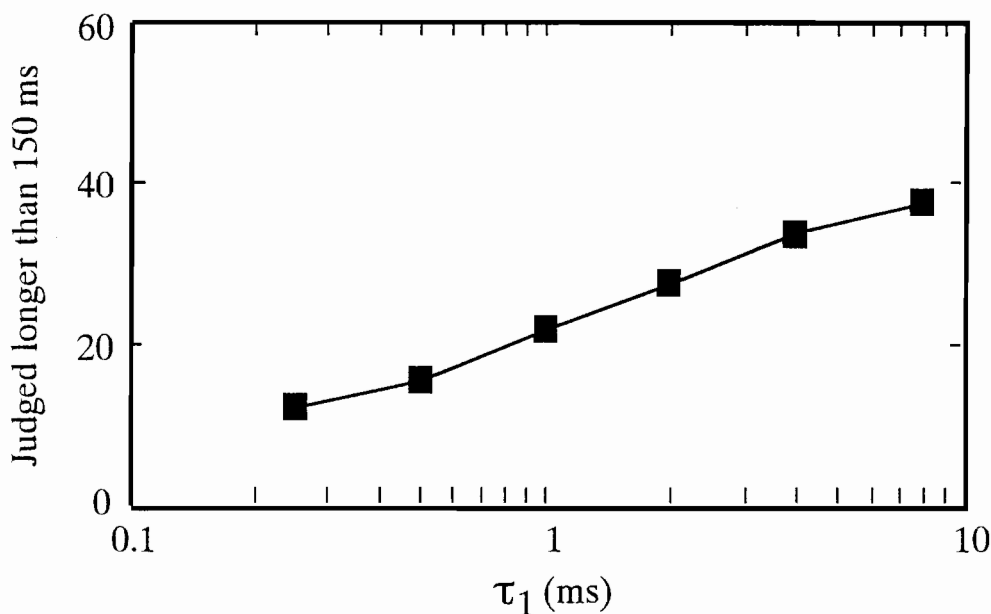
Figure 2.5. Durations were judged more longer for the lower frequencies or larger τ_1 of pure tones as standard stimuli (150 ms) than that of higher frequency or smaller τ_1. The comparison stimuli were the white noise. ■: 50% line of correct judgment meams DS.

2.3.2. Experiment 2

Average value of the duration sensation (DS), of the five responses at each pair for each subject (out of 10) was obtained. The cumulative frequencies of the correct judgments of the subjects for the three different stimuli presented first in the pair are shown in Figure 2.6. The reference physical duration was 150 ms, and the 50% line of correct judgments was defined to be the DS. The values on the 50% line of the cumulative frequencies of the averaged correct judgments are found to be 190.0, 185.0, and 180.0 ms respectively (longer than 150 ms are 26.7, 23.3 and 20 %, respectively), for the complex-tone stimuli whose fundamental frequencies were 250, 500, and 1000 Hz. The scores on the 50% line of cumulative frequencies of correct judgments with the individual differences (standard deviation = 7.37 ms) are shown in Figure 2.7. The

values of the analysis of variance (ANOVA) of the judged durations by ten subjects among the three stimuli are listed in Table 2.3.

Standard
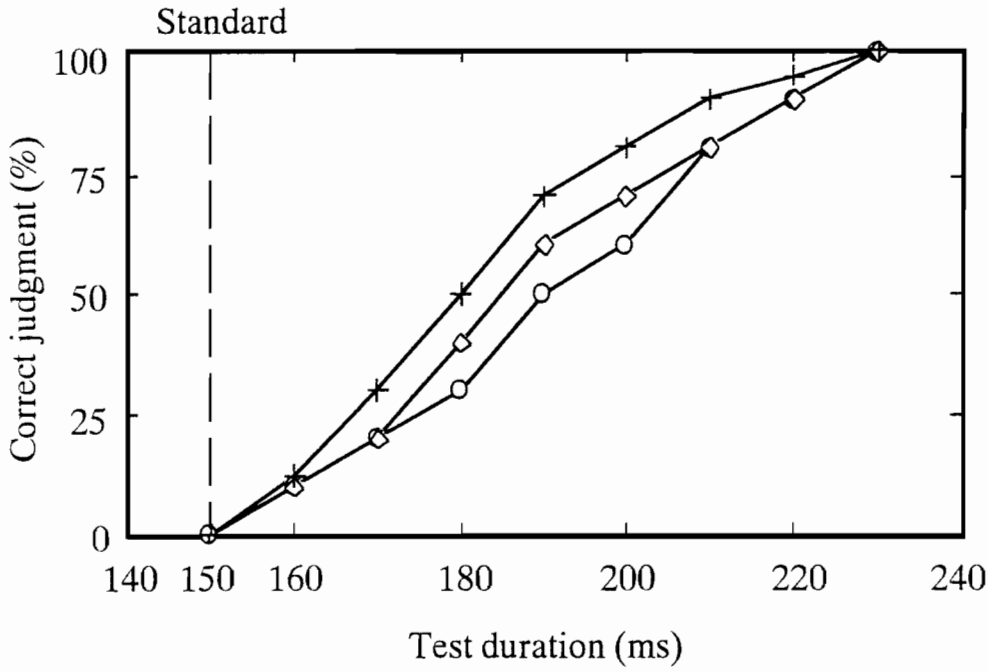
Figure 2.6. Cumulative frequencies of correct judgments of duration in the paired-comparison test for three stimuli. Stimuli were three complex tone with the different fundamental frequencies (f_0). Symbols signify the fundamental frequencies: $\bigcirc$: 250 Hz; $\diamondsuit$: 500 Hz; $+$: 1000 Hz. Dashed line indicates standard duration 150 ms.

Table 2.3. F-values of the analysis of varience for the judged-durations among the three stimuli of complex tone. f_0: fundamental frequency of the complex tone. Compared stimulus at second in pair was white noise.

f_0 (Hz) =	250	500	1000
f_0 (Hz) = 250	—	5.07*	9.73**
500		—	4.66
1000			—

* and ** indicate 5% and 1% significant level, respectively.
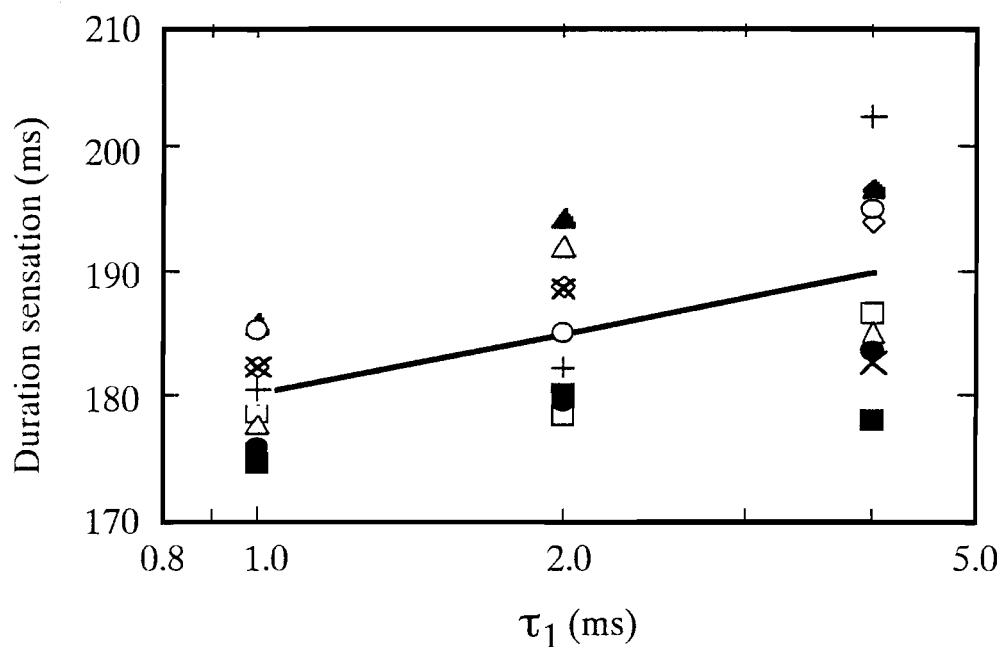
Figure 2.7. Judged durations for complex-tone stimuli with the three fundamental frequencies. The symbols indicate individual score of duration sensation obtained by the ten subjects (■: JA; □: TK; ◆: KF; ◇: RS; ▲: NA; △: TH; ●: KS; ○: NK; ×: YO; +: NA). ▬: average.

2.3.3. Experiment 3

Results from the experiment 3, indicate the comparison of the *DS* on the scale value (SV) obtained by using the "simple method of calculating individual subjective responses of paired-comparison test" (Ando and Singh; 1996) between pure-tones and complex-tone stimuli. The individual scale values of the judged durations for ten subjects are shown in the Figure 2.8. The average of those individual data are also shown in the Figure 2.9. The three graphs in the Figure 2.9 are the results between the pairs of pure-tone and pure-tone; complex-tone and complex-tone; and pure-tone and complex-tone stimuli. The pairs were compared with each other on the basis of three

stimulus-durations of 140, 150 and 160 ms.
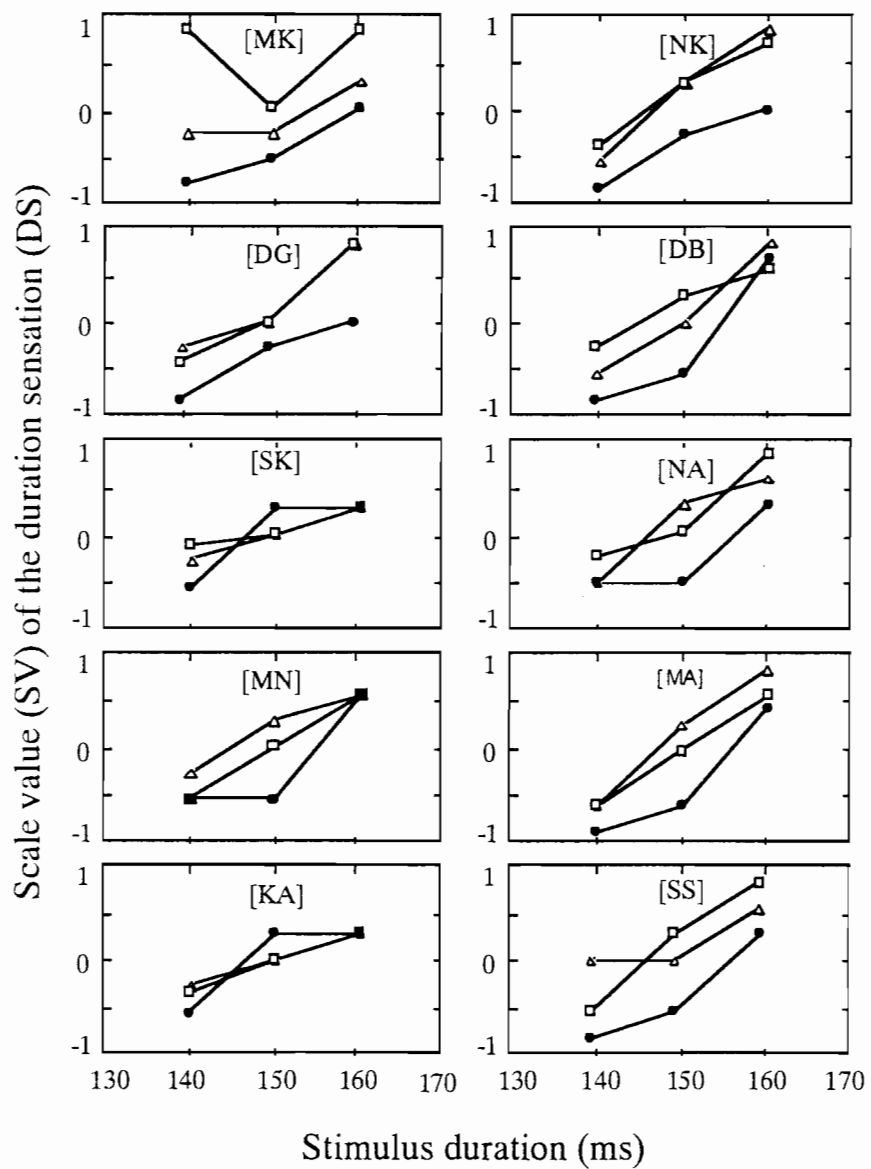
Figure 2.8. Individual scale values of the judged durations of 10 subjects (MK, DG, SK, MN, KA, NK, DB, NA, MA and SS) for the three stimulus durations. Symbols signify □: complex tone (500 Hz); △: pure tone (500 Hz); and ●: pure tone (3000 Hz).
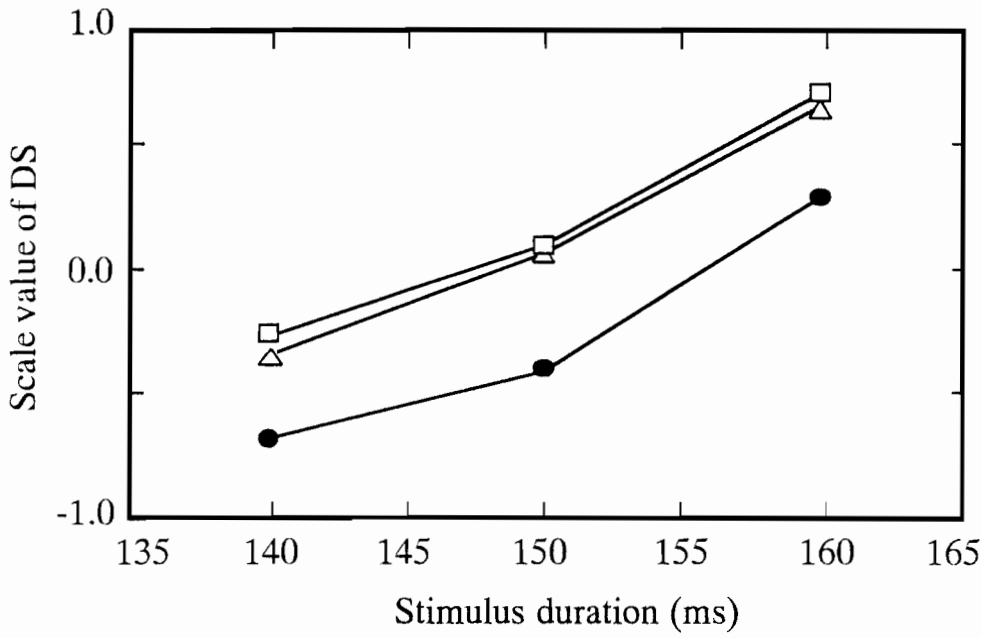
Figure 2.9. Scale values of duration judgments for three durations obtained by the paired-comparison test. $\square$: complex tone, fundamental frequency 500 Hz with f_6 and f_7 pure tone components; $\triangle$: pure tone 500 Hz; and $\bullet$: pure tone 3000 Hz.

The results (in Figure 2.9) show that the DSs were found almost similar with a non significant difference (significant level: $p > 0.05$) between pure tone and complex tone with the same values of τ_1 ($\tau_1 = 2$ ms) which correspond to the same pitch. But the DSs of pure tone (frequency: 3000 Hz; $\tau_1 = 0.33$ ms) were found significantly different ($p < 0.01$) with the pure tone (frequency: 500 Hz; $\tau_1 = 2$ ms) and complex tone (fundamental frequency: 500 Hz; $\tau_1 = 2$ ms) both stimuli.

2.4. Discussion

2.4.1. Experiment 1

In Figure 2.3, the results with a shorter duration indicate that the first stimulus is judged to be shorter (higher frequency pure tone) and the second stimulus is judged as longer (white noise) in the presented pair. Judging the second stimulus in the pair as longer (except the stimuli of 140 and 150 ms) makes the correct response (according to instructions) plotted on the graph and appear to be of shorter duration. The graph means that the judged duration was longer for a pure tone with a lower frequency (larger τ_1) in the Figure 2.5. The 50% line of correct judgments was defined here as the duration sensation. Frequency was thus found to have an effect on the human sensation of duration in that pure tones with higher frequencies were judged to be shorter than pure tones with lower frequencies when their durations were compared with the duration of a white-noise test stimulus. According to Figure 2.3, duration sensations (DS) on the 50% line of judgments are dependent on the value of physical duration (PD) of the used stimuli.

2.4.2. Experiment 2

The judged-durations for each individual and the average for the durations of the three stimuli of complex-tone were found to be significantly longer ($p < 0.01$) between the complex tones with the fundamental frequencies of 250 and 1000 Hz as shown in the Table 2.3. Figure 2.6 also shows the DSs are found different when, physical durations (PD) of the stimuli are different.

Table 2.4. Values of 't' for the judged-durations in the paired-comparison test, among the stimuli of one complex tone and two pure tones with three durations (ms). Fundamental frequency of 500 Hz complex tone was made by the f_6 and f_7 components pure tone.

	PT - 500Hz			CT - 500Hz			PT - 3000Hz		
	140	150	160	140	150	160	140	150	160
PT-500Hz									
140[ms]	—	6.7**	9.0**	0.7	7.8**	5.1**	2.8**	8.1**	13.4**
150		—	3.8**	-6.3**	-0.3	1.8	-2.3*	4.3**	7.4**
160			—	-13.1**	-7.2**	-0.6	-4.3**	0.6	4.1**
CT-500Hz									
140[ms]				—	11.0**	7.3**	2.6*	8.5**	15.0**
150					—	2.4*	-1.4	5.0**	11.2**
160						—	-2.7*	0.9	2.8*
PT-3000Hz									
140[ms]							—	4.4**	7.5**
150								—	2.7*
160									—

* and ** indicate 5% and 1% significant level, respectively.

Table 2.5. Values of analysis of variance (ANOVA) for the judged-durations in the paired-comparison test, among the stimuli of one complex tone and two pure tones with three durations (ms). Fundamental frequency of 500 Hz complex tone was made by the f_6 and f_7 components pure tone.

	PT - 500Hz			CT - 500Hz			PT - 3000Hz		
	140	150	160	140	150	160	140	150	160
PT-500Hz									
140[ms]	—	0.6**	1.0**	-0.1	0.6**	0.9**	0.4**	1.1**	1.3**
150		—	0.4**	-0.6**	-0.03	0.3**	-0.2	0.5**	7.4**
160			—	-1.1**	-0.5**	-0.1	-0.6**	0.1	0.3**
CT-500Hz									
140[ms]				—	0.6**	1.0**	0.4**	1.1**	1.4**
150					—	0.4**	-0.2	0.5**	0.8**
160						—	-0.6**	0.1	0.4**
PT-3000Hz									
140[ms]							—	0.7**	1.0**
150								—	0.3*
160									—

* and ** indicate 5% and 1% significant level, respectively.

2.4.3. Experiment 3

In the experiment 3, the results indicate that the duration sensations are found almost similar for the pure-tone ($\tau_1 = 2$ ms) and complex-tone ($\tau_1 = 2$ ms) stimuli. But the DSs of the pure tone ($\tau_1 = 0.33$ ms) were found significantly (p $<$ 0.05) different from the pure-tone ($\tau_1 = 2$ ms) and complex-tone ($\tau_1 = 2$ ms) stimuli. It means the τ_1 in the ACF of the sound stimulus is responsible for the duration sensation (DS). In the Figure 2.9, the distances of scale values between the two pure tones ($\tau_1 = 2$ ms and 0.33 ms) for the three stimulus durations were found almost same (0.34, 0.36, 0.36). It means that the effect of τ_1 on the DS is independent for the stimulus durations.

Table 2.6. Results of the tests of goodness of fit for the ten subjects in which, k: number of defective pairs; d: percentage of defective pairs; λ : ratio; and σ : standard deviation.

Subject	k	d	λ	σ
MT	8	22.2	0.21	0.57
FK	6	16.7	0.02	0.54
SK	6	16.7	0.05	0.31
MK	4	11.1	0.09	0.50
NK	7	19.4	0.13	0.52
DB	6	16.7	0.15	0.61
RA	4	11.1	0.06	0.62
NK	5	13.9	0.04	0.52
DB	5	13.9	0.13	0.65
RA	6	16.7	0.03	0.56

To justify the differences of the scale values of *DS* among the three stimuli with three stimulus-durations 't-test' was used (see Table 2.4). The values of the analysis of variance (ANOVA) also shown in the Table 2.5. A 'test of goodness of fit' was also applied to signify the data which, is shown in Table 2.6.
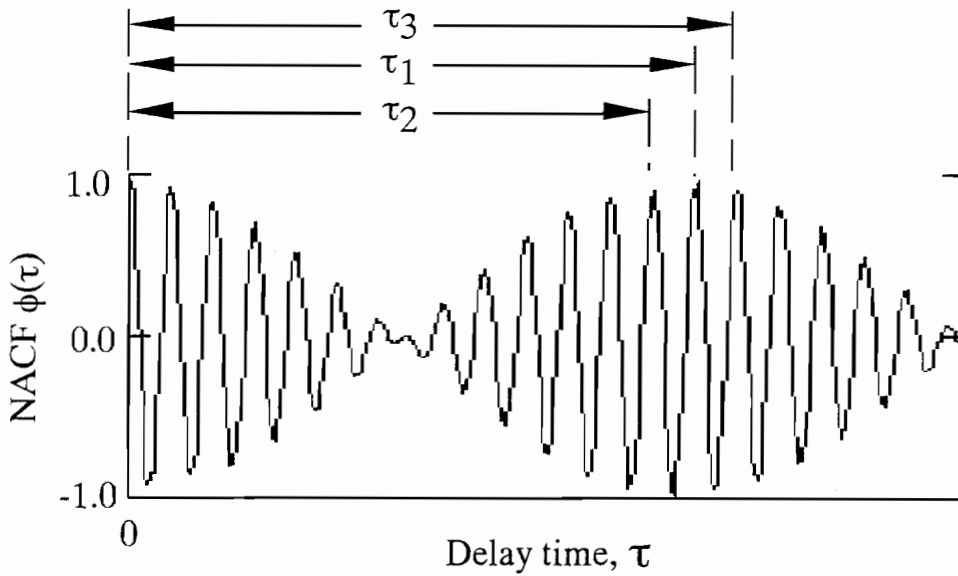
Figure 2.10. An example of the wave-form of normalized autocorrelation function (NACF) of a complex-tone stimulus. The τ_1 is the time delay of the first peak (highest one) correspond to the fundamental frequency. The τ_2, τ_3, and contemporary are associate peaks extracted from the NACF.

For the complex-tone stimuli (see Figure 2.10) τ_2, τ_3, . . . τ_n can not make interference (Schouten, 1940) to determine the *DS*. Because τ_2 and τ_3 has the inverse direction from the τ_1. So the effect of one associate peak is canceled by the another oposite one.

So it can be formulated that the τ_1 of ACF of pure-tone and complex-tone stimuli are the function of DS. So that in the results of the three experiments, DS can be recognized as the fourth of the four primary sensations of sound stimuli after -loudness, pitch, and timbre. The DS also found as dependent on the physical duration (PD) of the stimuli (Figures 2.3; 2.6; 2.9). So DS is expressed here by

$$DS = f(\tau_1, PD), \tag{2.3}$$

Here smaller PD means minor effect on DS.

2.5. Conclusions

The significant results of the study are cited briefly below.

1. The DS of pure-tone stimuli with larger τ_1 (lower frequency) are judged as longer than those with the smaller τ_1 (experiment 1).

2. The duration sensation (DS) of the complex-tone stimuli with lower fundamental frequencies or lower pitch (larger τ_1) was judged to be longer than those stimuli with higher fundamental frequencies or higher pitches (experiment 2).

3. Longer DSs were obtained for the larger physical duration (PD) of the stimuli meaning that the PD is the function of DS (experiments 1 and 2).

4. The effects of τ_1 extracted from the ACF on the DS were found almost same on the scale value for the pure-tone ($\tau_1 = 2$ ms) and complex-tone ($\tau_1 = 2$ ms) stimuli that is the same pitch of the frequency and fundamental frequency (experiment 3).

5. The DS of the pure-tone stimulus ($\tau_1 = 0.33$ ms) was found significantly different

from the pure-tone and complex-tone stimuli with the different pitch ($\tau_1 = 2$ ms) (experiment 3).

6. The effect of the τ_1 on the DS also found as independent factor for the stimuli-durations (experiment 3).

CHAPTER 3

DS IN RELATION TO THE FACTOR ϕ_m $(\phi_m = \phi_1; m = 2,3,...)$ OF ACF FOR THE MIXED SITIMULI OF PURE-TONE AND WHITE-NOISE

3.1. Preface

The human experience of the passing of time duration has been studied from various points of view. Gyuau (1902) found that estimates of time depend on many variables, such as the number, type, and intensity of sound stimuli presented and the subject's expectations, level of attention, and interest. Woodrow (1951) described his data in terms of "Vierordt's law": that shorter time intervals are overestimated and longer intervals are underestimated. He also described an indifferent range in which intervals are neither overestimated nor underestimated. Frankenhaeuser (1959) presented clicks as acoustical stimuli, at various intervals ranging from 13 to 72 s and sequenced at various speeds, and found that estimates of duration increased with the speed at which the clicks were sequenced. Fraisse (1961) found that an interval during which signals are presented at five per second is estimated to be longer than an interval in which they are presented at ten per second. Michon (1965) described the variation of subjective duration that was due to variations in the types and amounts of activity in which the subjects were engaged. Austin and Marilyn (1966) used auditory clicks at frequencies between zero and ten per second and found that the magnitude of the

frequency effect was a monotonically decreasing function of duration. Vorn (1970) found that the apparent durations increase with the rate of information transmission and that when a subject has to behave actively in this case, by translating the stimuli into binary choices the experienced duration decreases with the number of processed bits. Ando (1977) found that time passes more quickly when it is filled with noise than when it is silent and explained this to an internal clock inhibited by noise. Zakay et al. (1983) found that subjective estimates of time were a decreasing function of task difficulty, and that estimates of the durations of "empty" intervals were longer than that of "filled" intervals. Subjects produced the longest estimates of duration when the external tempo was fastest and produced shorter estimates when the external tempo were. Pastor and Artieda (1996) stated that the internal clock can be imagined to be an oscillating circuit, with a single pacemaker neuron or a group of pacemaker neurons, that generates rhythmical electromagnetic activity the brain uses in the temporal analysis of information.

Given this background, it can be thought that the effect of environmental noise on the duration sensation should be studied. The work reported in this paper, environmental noise was simulated by mixing pure-tone and white-noise. Such a mixture is perceived to be similar to the noise of an aircraft landing. Environmental noise simulated this way can be easily characterized in terms of the ratio of the two sound pressure levels (SPLs) and the frequency of the pure tone.

This study is an investigation of the ways that the duration sensation (DS) of simulated environmental noise depends on certain factors extracted from the ACF of

that noise ($\Phi(0)$): energy at the origin of sound stimulus; τ_1: time delay of the first peak; ϕ_m: amplitude of peaks, 'at the m-th peaks where' m = 1, 2, 3, . . . , n). Two experiments (Experiment 1 and 2) were conducted with the inverse way of subjective judgments in the paired-comparison method. The ϕ_m was varied by changing the ratio of sound pressure levels (SPLs) as $\Phi(0)$ extracted from the ACF of the mixed state of the pure-tone and white-noise stimuli.

3.2. Methods

3.2.1. Stimuli preparing

Five sound stimuli were selected to use in the paired-comparison test for the subjective judgment of its duration. The white noise, one pure-tone stimulus (1000 Hz) and more three mixed stimuli were included in each (experiments 1 and 2) experiment. The two components of three mixed stimuli were pure tone and white noise with the difference of their SPL. The pure-tone stimulus contained a constant SPL, 80 dB(A) whereas the SPL of white-noise stimulus were changed. Two stimuli were presented at a time by the two loudspeakers in the sound proof chamber. The rise and fall times of all stimuli in both experiments were 1 ms. The times were defined as the time taken to reach a level −3 dB different from the steady level.
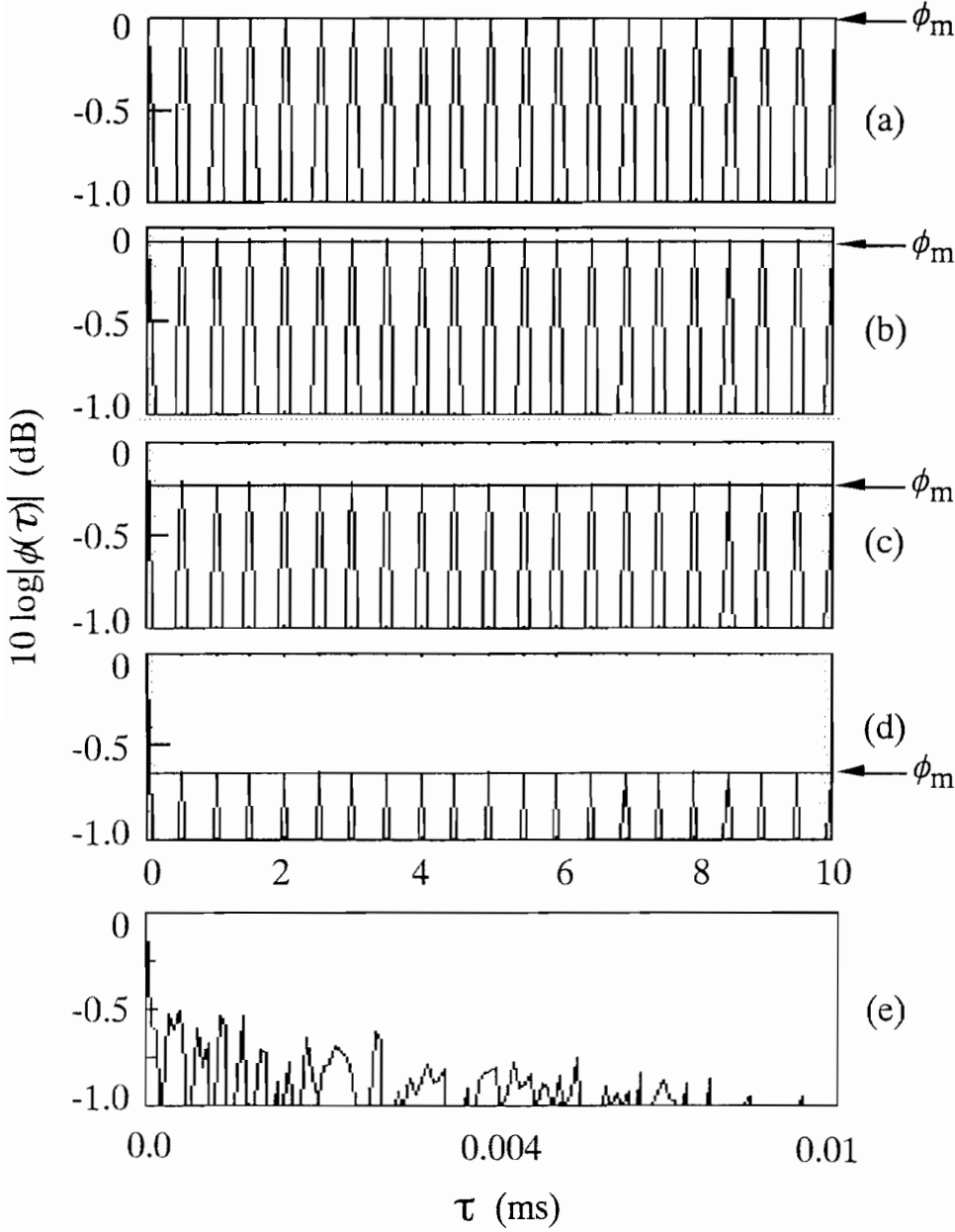
Figure 3.1. Measured ϕ_m (m = 1, 2, 3, . . ., n) of the ACF envelope for four sound stimuli. (a) pure tone (P) 1000 Hz, 80 dB (A), $\phi_m = 0.0$ dB; (b) pure tone plus 59 dB(A) white noise (PW1), $\phi_m = -0.06$ dB; (c) pure tone plus 64 dB(A) white noise (PW2), $\phi_m = -0.23$ dB; (d) pure tone plus 69 dB(A) white noise (PW3), $\phi_m = -0.64$ dB. Note that the measured ACF of white noise was obtained (e), where $\phi(0) = 1$ and $10\log\phi(\tau) < -10$ dB, $\tau > 0$.
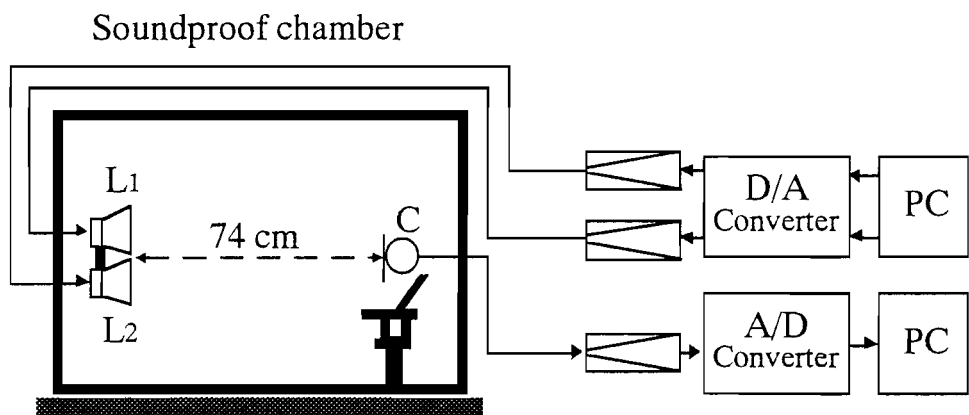
3.2.2. ACF measurement

The values of the peak amplitudes (ϕ_m) of the ACF envelope were extracted from the short-term moving autocorrelation function of each of five stimuli (five seconds each) over a constant integration interval of 2T = 2 s, as shown in Figure 3.1. The stimuli were recorded in a soundproof chamber through an A-weighted filter (Figure 3.2). Note that the changes of the value of 2T have no effect on ϕ_m in this investigation. Such values of ϕ_m are extracted as constant. This means that the envelope of the ACF on a logarithmic scale is almost parallel to the horizontal axis (ϕ_m = 0 dB line for pure tones). For white noise however, the envelope of ϕ_m values is not parallel to the zero level. Values were measured for an 80 dB(A) pure tone: ϕ_m = 0.00 dB; an 80 dB(A) pure tone plus 59 dB(A) white noise: ϕ_m = -0.06 dB; an 80 dB(A) pure tone plus 64 dB(A) white noise: ϕ_m = -0.23 dB; an 80 dB(A) pure tone plus 69 dB(A) white noise: ϕ_m = -0.64 dB; and for 80 dB(A) white noise alone (Table 3.1).

Table 3.1. Measured and calculated values of the peak amplitudes (ϕ_m) of the envelope of ACF for five stimuli.

Stimulus	Measured values of 10 log| ϕ_m| dB	Calculated values of 10 log| ϕ_m| dB
P: PT 80 dB(A)	0.00	0.00
PW1: PT 80 dB(A) plus WN 59 dB(A)	-0.06	-0.08
PW2: PT 80 dB(A) plus WN 64 dB(A)	-0.23	-0.25
PW3: PT 80 dB(A) plus WN 69 dB(A)	-0.64	-0.75
W: WN 80 dB(A)	($<$-10.00)	$-\infty$

The values in the parentheses is estimated.

For the same set of stimuli the calculated peak amplitudes of the ACF envelopes were respectively 0.00 dB, -0.08 dB, -0.25 dB, and -0.75 dB, (These values were calculated as described in the APPENDIX - B). The values of $\phi(\tau)$, given by equation (B.3) in the APPENDIX - B, were respectively 0.99, 0.97, and 0.92 for three mixed stimuli in which $\phi_m = 10\log_{10}\phi(\tau)$. The measured and calculated values of ϕ_m for the first four of these stimuli (i.e., excluding the value for the white-noise stimulus) are plotted in Figure B.a of the APPENDIX - B and can be seen to be very close to a straight line.

Soundproof chamber

Figure 3.2. Block diagram of the stimuli presentation for the ACF measurement and subjective test. L: loudspeaker; C: center of the microphone and human head; D/A: digital to analoge; A/D: analoge to digital; PC: personal computer.

3.2.3. Stimuli presentation

The first (standard) stimulus in these tests was a pure tone (1000 Hz) with a duration of 150 ms and a SPL of 80 dB(A) for the both experiments. Ten sets of stimuli with a

ten-ms step size in the range from 140 to 230 ms were used in each session of the experiment 1. Ten sets of stimuli changing from 70 to 160 ms (with the same step size) were presented in each session of the experiment 2. The five stimuli in each set had five different ϕ_m values for their ACF envelopes. The SPL of all stimuli was kept at a constant 80 dB(A). In the paired-comparison test, the second stimuli were presented in random order. Each pair was presented 20 times repeatedly. The intra pair and inter pair (response time) gaps respectively were 1 and 3 s.
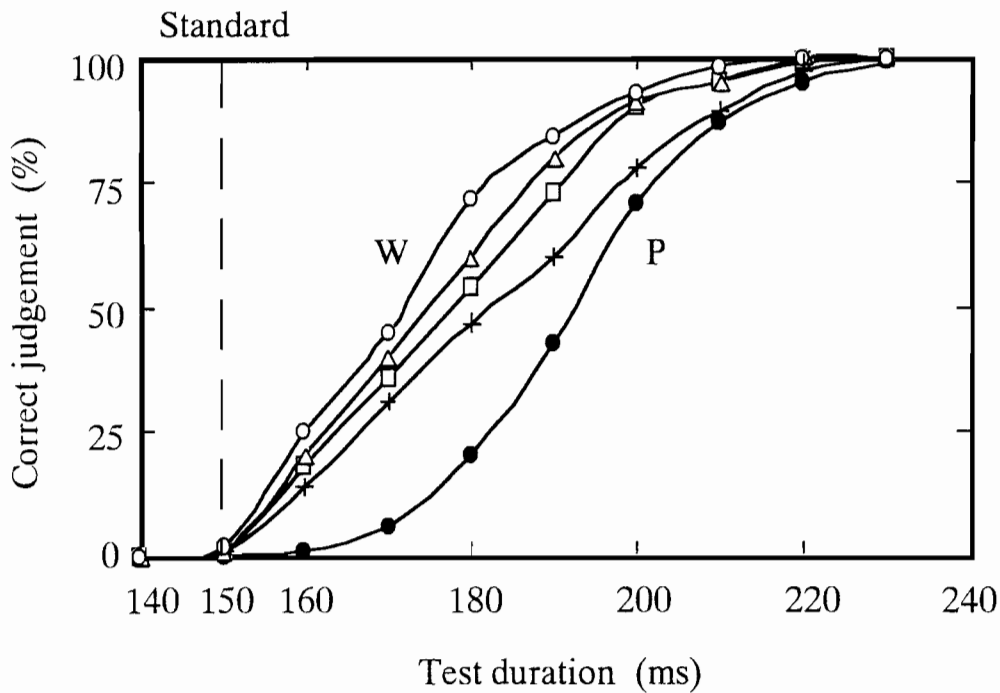
Figure 3.3. Cumulative frequencies of correct judgements indicate the 50% line of durations for five stimuli: ●(P), ϕ_m = 0.00 dB; + (PW1), ϕ_m = −0.06 dB; □ (PW2), ϕ_m = −0.23 dB; △ (PW3), ϕ_m = −0.64 dB and ○ (W), white-noise 80 dB(A). The dashed line indicates physical duration of the standard stimulus was 150 ms. Results of first test: judging whether the duration of the test stimuli is longer than that of the standard stimulus.

3.2.4. Subjective test

The subjects of both experiments were five male (same) students in the laboratory (KF, MI, NS, OK and YS). All of them have good hearing and were between 22 and 25 years old. The subjects were seated in the soundproof chamber on a chair in front of two loudspeakers aligned vertically with their centers 15 cm apart (Figure 3.2). One loudspeaker provided the pure tone and the other provided the white noise. The distance between the center of the subject's head and the front of the loudspeakers was 74 ($\pm$1) cm during subjective test.
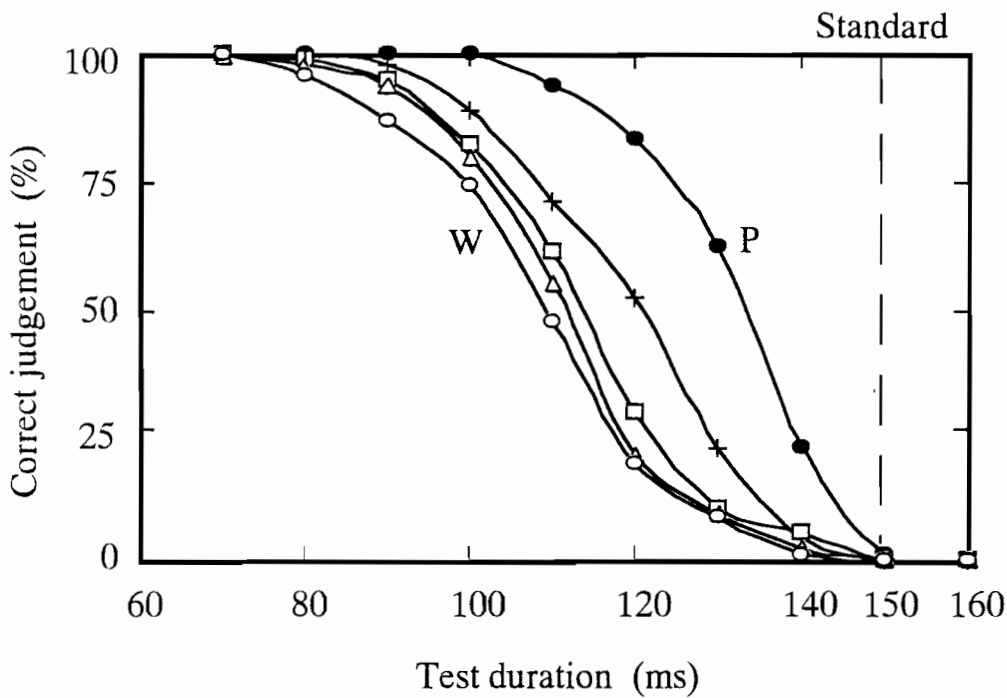
Figure 3.4. Cumulative frequencies of correct judgements indicate the 50% line of durations of five stimuli ●(P), ϕ_m = 0.00 dB; + (PW1), ϕ_m = -0.06 dB; □ (PW2), ϕ_m = -0.23 dB; △ (PW3), ϕ_m = -0.64 dB and ○(W): white-noise 80 dBA. The dashed line indicates physical duration of the standard stimulus was 150 ms. Results of second test: judging whether the duration of the test stimuli was shorter than that of the standard stimulus.

A constant standard stimulus (pure tone, 150 ms) was used which, consisted of two experiments (experiments 1 and 2). Most of test stimuli in the experiment 1 were slightly longer than the standard stimulus and most of those in the experiment 2 were slightly shorter than the standard stimulus. The subjects were asked to judge whether the duration of the second of a pair of stimuli (the test stimuli) was longer than or shorter than the first standard stimulus. In the experiment 1, subjects were told to push the button only when the second stimulus seemed longer than the standard. In the experiment 2, subjects were told to push the button only when the second stimulus seemed shorter than the standard. Each subject went through five sessions with five different test stimuli for each test.

Table 3.2. Values of judged durations (ms) obtained by the paired-comparison test in respect to five sound stimuli (Figures 3.3 and 3.4). PW1, PW2 and PW3 are describe in the Table 3.1.

Lines of percentage [%]	(P) Pure tone	PW1	PW2	PW3	(W) White noise
	Experiment 1				
50% line	193.0 28.7*	182.5 21.7*	177.5 18.3*	175.0 16.7*	172.0 14.7*
	Experiment 2				
50% line	134.5 9.7*	121.5 19.0*	115.0 23.3*	112.0 25.3*	109.5 27.0*

*: percentage of the deviation in duration than 150 ms.

3.3. Results

Cumulative frequencies of the correct judgment of duration sensation for the five subjects were obtained by the effect of five different stimuli as a parameter of different values of ϕ_ms in the ACF, as shown in Figures 3.3 and 3.4. In the experiment 1 and 2, values of correct judgements were obtained as duration cross the 50% line for the stimuli of P, PW1, PW2, PW3, and W, respectively in the Table 3.2. A significant difference ($p < 0.05$) was found between results of the five test stimuli of both experiments for each 50% crossing lines. The results of the analysis of variance (ANOVA) of the judged durations for all subjects are listed in Table 3.3.
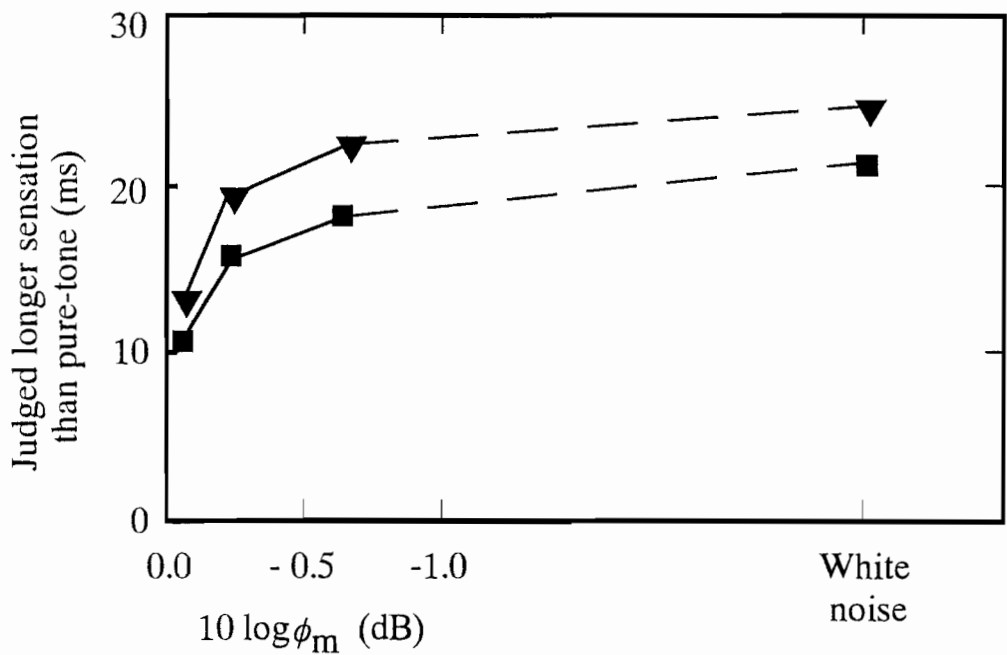
Figure 3.5. Longer sensations of the judged durations for four stimuli including white-noise than the judged reference of pure-tone stimulus. Duration sensation were defined by the 50% line of correct judgement. ■: experiment 1 and ▼: experiment 2.

Taking the results (●) for the pure tone in Figures 3.3 and 3.4 as a judged reference, other four of the results ($+$, □, △ and ○) show a significant ($p < 0.01$) deviation in the judgment of durations. The deviations of duration from the result of the judged reference (pure tone) for the experiments 1 and 2 are shown together in the Figure 3.5.

Table 3.3. F-values of the analysis of varience for the judged-durations with five subjects among the five stimuli. The meaning of the P, PW1, PW2, PW3 and W are described in Table 3.1. Values without and within bracket indicate experiment 1 and 2, respectively.

	P	PW1	PW2	PW3	W
P	—	10.4 (14.7**)	17.3* (17.3**)	19.1* (25.8**)	23.5** (30.0**)
PW1		—	6.9 (2.6)	8.7 (11.1*)	13.1 (15.3**)
PW2			—	1.8 (8.5*)	6.2 (12.7**)
PW3				—	4.4 (4.2)
W					—

* and ** indicate 5% and 1% significant level, respectively.

3.4. Discussion

In both experiments, all five test stimuli produced different results when subjects compared them with the one standard stimulus. The 50% line of correct judgments was defined as duration sensation (DS). The results indicate a tendency for smaller attenuation levels (ϕ_m) of the envelope of the ACF for the test stimuli to produce a

longer DS for both experiments (Figures 3.3 and 3.4). A lower ϕ_m, however, produces a shorter difference from the physical standard stimulus (150 ms; dotted lines in the Figures 3.3 and 3.4) in the experiment 1 and an inversely longer distance in the experiment 2. This is because of the nature of the judgments in the two experiments: whether longer-judgment for the longer test stimuli in the experiment 1, and shorter-judgment for the shorter test stimuli in the experiment 2. Shorter durations on the graph thus indicate a longer DS in comparison with the judged reference pure tone (P: in the Figures 3.3 and 3.4).

In Figure 3.5, data for white noise (the ϕ_m cannot be graphed exactly) can be explained as a continuation of data for the other three values of ϕ_m. It is clear, however that such envelope of ACF may be an indicator of the DS. So the effects of the ϕ_m of the envelope of an ACF for sound signals was found to have a significant effect on the human duration sensation (DS). The both experiments show the DS is to be different if the physical duration (PD) of the sound stimuli is different. So the DS can be found as a function of the envelope of ACF and physical duration according to the below equation.

$$DS = f(\phi_m, PD), \qquad\qquad (3.1)$$

here, smaller PD means minor effects on the DS.

3.5. Conclusions

From the results of this study (experiments 1 and 2) it can be concluded that

1. The duration sensation (DS) of a white-noise stimulus is judged to be longer than that of pure-tone stimuli (125–4000 Hz) under the same sound pressure level. That actually has the same duration.

2. The DS of mixed stimuli (pure tone and white noise) with higher values of attenuation level (larger ϕ_m) of the envelope of ACF are judged as shorter sensation than that of the smaller ϕ_m.

3. Physical duration (PD) has the effects on the DS.

CHAPTER 4

DS IN RELATION TO THE FACTORS τ_1, ϕ_1 AND τ_e OF THE ACF OF BANDPASS-NOISE STIMULI

4.1. Preface

A theory of the primary sensation of environmental noise -loudness, pitch and timbre- is described by Ando, 2001a. The sensation of environmental noise can be well described by factors extracted from the autocorrelation function (Ando, 1998). One study was conducted on the duration sensation as the fourth primary sensation of simulated environmental noise (Saifuddin et al. 2001a).

In the previous study Saifuddin et al. (2001a) investigated the duration sensation for stimuli consisting of the white noise with different sound pressure level (SPL) and a pure tone. Thus, the attenuation level of the envelope (ϕ_m) of the autocorrelation function (ACF) of the stimuli was changed the amplitude difference between the white noise and pure tone. It was found that the duration of stimuli with smaller ϕ_m was judged as being longer sensation than that of larger ϕ_m. In the same study, the judgment of longer-duration of sensation for the pure tone was found when the frequency is lower (larger τ_1) of the tone. Except for the study on auditory gap detection by Oxenham (2000) the effects of bandpass noise on the duration sensation has not been studied before. In his study, a larger deterioration was observed when two markers of

broadband noise occupied different spectral regions but had the same fundamental frequency.

The loudness of bandpass noise has been investigated in relation to the center frequency and the bandwidth of the stimulus. In the classical theory of bandpass noise, the loudness of bandpass noise remains constant as the bandwidth of the noise increases until the noise reaches the "critical band", after which the loudness increases with further increases of bandwidth (Zwicker et al. 1957 and 1972, Greenwood 1961, and Hellman 1970). Merthayasa et al. found that the loudness of sharply filtered noise centered on 1000 Hz (1080 dB/oct.) within the critical band is not constant (Merthayasa et al. 1994 and 1996). Ando (1998) showed that loudness is influenced by the increasing values of effective duration (τ_e) of ACF, which is defined by the delay at which the envelope of the normalized ACF becomes –10 dB. Florentine et al. (1996) compared the loudness of 1000 Hz tones with that of broadband noise over a wide range of levels. They found that the amount of temporal integration, defined as the level difference between equally louder stimuli with the duration of 5 and 200 ms, varies with level.

Fujii et al. (2001) found that flying aircraft-noise, which is a mixed state between 'tonal noise' and 'un-tonal noise', can be well represented by the factors extracted from the ACF and the interaural cross-correlation function. Such aircraft noise as environmental one may be perceived as like as bandpass noise at the lower center frequency. Regarding these, to examine the effects of center frequency and the bandwidth of bandpass noises on the duration sensation (DS) can be meaningful if it may explain well by the factors extracted from the ACF. In this aspect the DS of

bandpass noises with different center frequencies and bandwidths including critical

bandwidth was studied here.
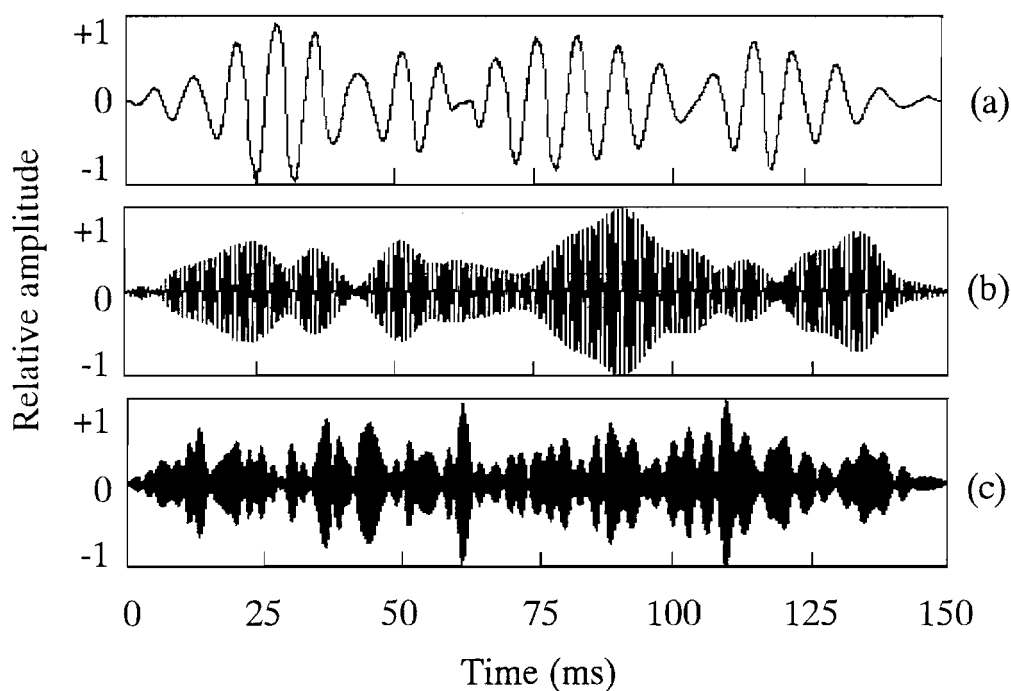
Figure 4.1. Examples of stimuli used in the study. (a) Center frequency f_c = 125 Hz and bandwidth Δf = 80 Hz; (b) f_c = 1000 Hz and Δf = 160Hz; and (c) f_c = 4000 Hz and Δf = 640 Hz.

4.2. Methods

4.2.1. Stimuli preparing

The white noise and twenty-two bandpass noises with combinations of six center

frequencies (125, 250, 500, 1000, 2000 and 4000 Hz) and different bandwidths were

used as stimuli. The bandwidth was changed with a cut-off slope of 2068 dB/octave,

which is obtained by the combination of two digital filters. The bandwidth of 0 Hz
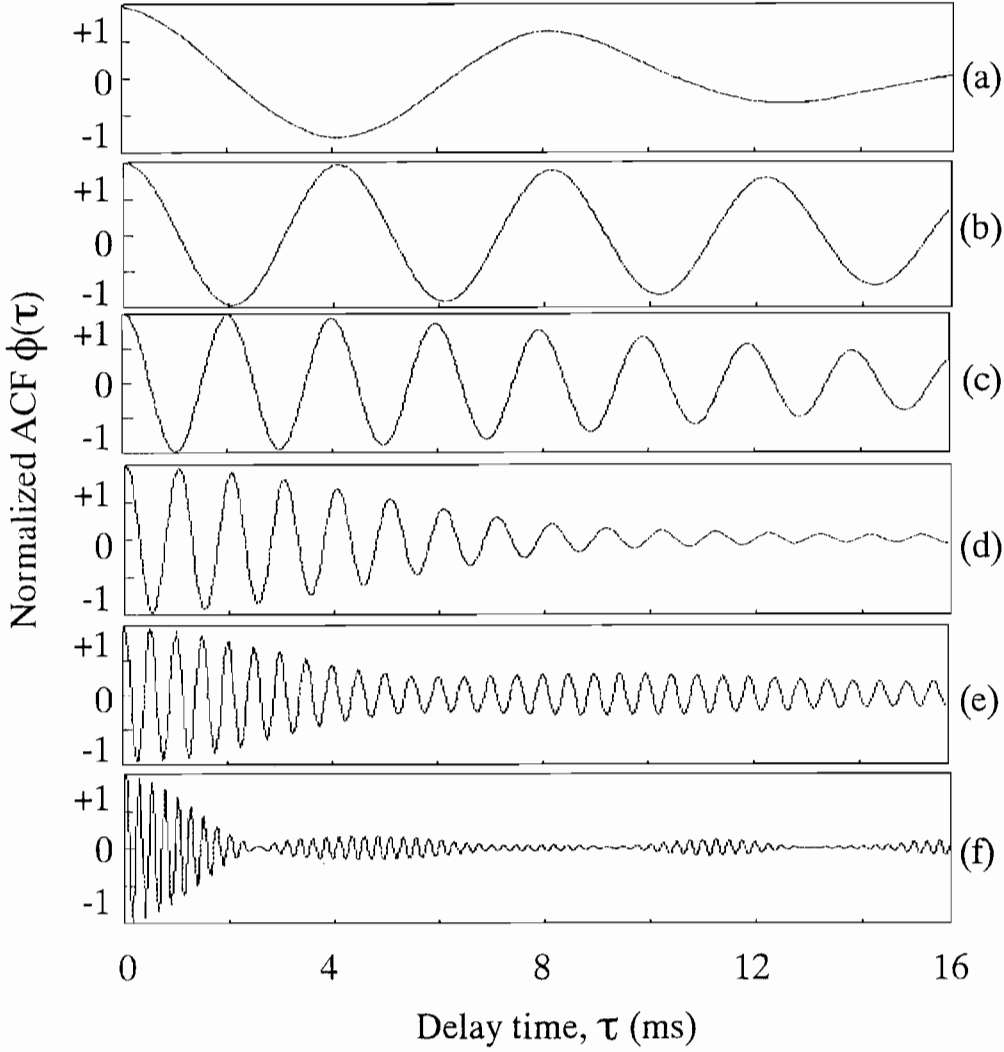
Figure 4.2. Measured normalized ACF of the six bandpass-noises (at the center position of subject's head in the listening room) with different center frequencies (f_c) for selected bandwidths (Δf). (a) $f_c = 125$ Hz and $\Delta f = 80$ Hz; (b) $f_c = 250$ Hz and $\Delta f = 80$ Hz; (c) $f_c = 500$ Hz and $\Delta f = 80$ Hz; (d) $f_c = 1000$ Hz and $\Delta f = 160$ Hz; (e) $f_c = 2000$ Hz and $\Delta f = 320$ Hz; and (f) $f_c = 4000$ Hz and $\Delta f = 640$ Hz.

means only the slope components. The rise and fall times of all stimuli were 1 ms. The

rise and fall times were defined as the time taken to reach a level −3 dB different from

steady level. Examples of stimuli used in the study are shown in Figure 4.1.
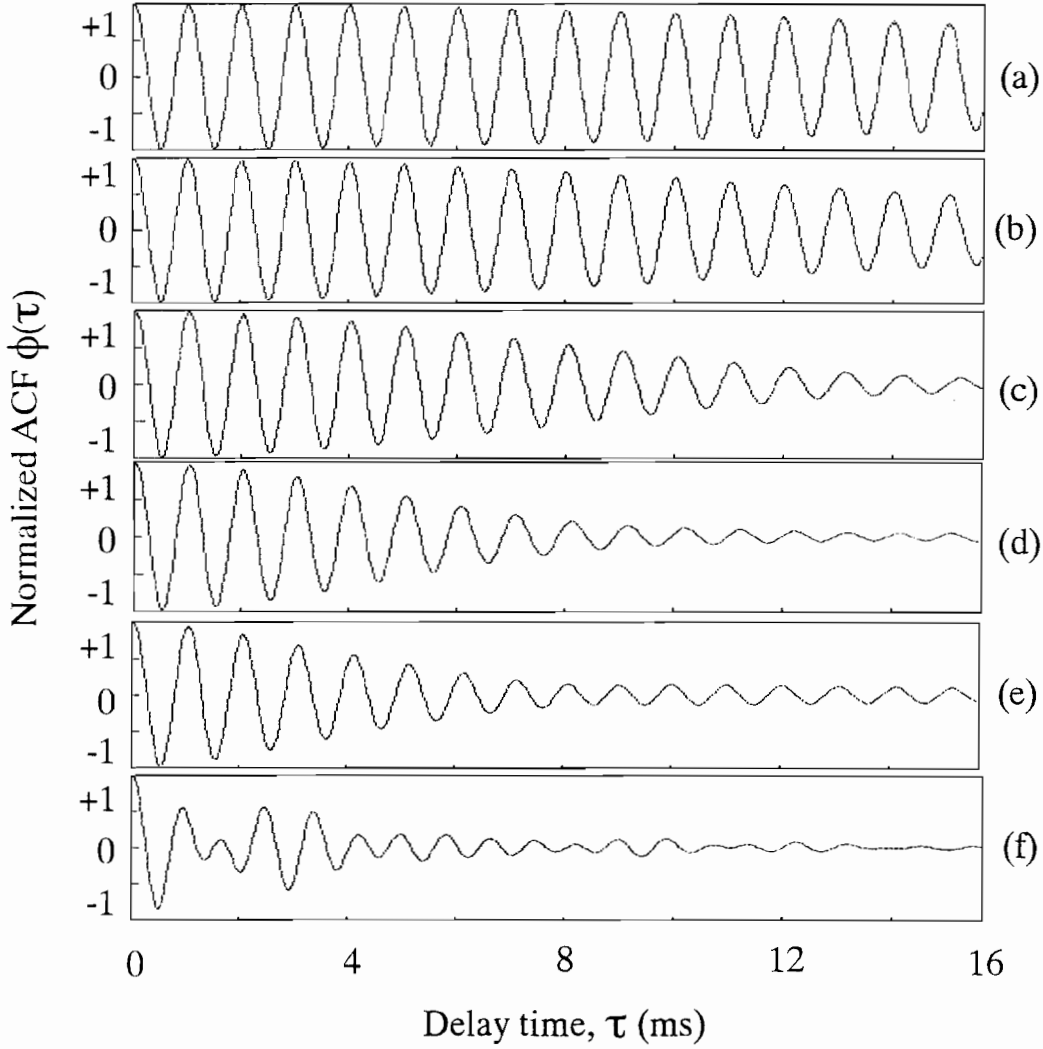
Figure 4.3. Measured normalized ACF of the six bandpass-noises with different bandwidths (Δf) (at the center position of subject's head in the listening room) whose the center frequency is 1000 Hz. (a) Δf = 0 Hz; (b) Δf = 40 Hz; (c) Δf = 80 Hz; (d) Δf = 160 Hz; (e) Δf = 320 Hz; and (f) Δf = 640 Hz.

4.2.2. ACF measurement

Stimuli were characterized in terms of their ACF. All twenty-two bandpass-noise stimuli were reproduced by a loudspeaker placed inside soundproof chamber and were recorded by a microphone at the horizontal distance of 74 ± 1 cm from the loudspeaker. The sound pressure level (SPL) of all stimuli were kept constant at 80 dB(A) accurately by measurement of the ACF at the origin of the time delay $\Phi(0)$. The normalized ACF of the stimuli after passing through an A-weighted network was calculated (the integration interval of ACF; 2T was 2s). Examples of the normalized ACF of the six center frequencies at the selected bandwidths are shown in Figure 4.2. Figure 4.3 show examples of the normalized ACF of the six bandwidths centered on 1000 Hz.
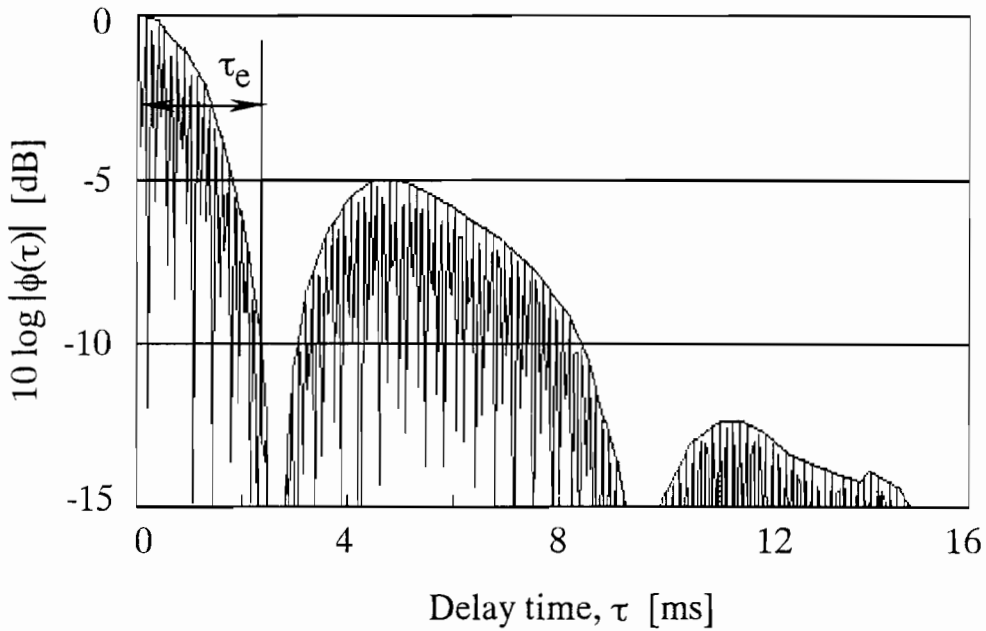
Figure 4.4. Example of determining τ_e of the bandpass-noise stimulus. The τ_e is defined by the delay time at which the first envelope of the autocorrelation function becomes -10 dB.

Table 4.1. Measured values of τ_1, τ_e and ϕ_1 extracted from the ACF of six bandpass-noise stimuli for different center frequencies (f_c) with selected bandwidths (Δf).

	τ_1 (ms)	τ_e (ms)	ϕ_1
f_c (Hz) = 125 Δf (Hz) = 80	7.40	108	0.65
250 80	4.00	52	0.94
500 80	2.00	25	0.96
1000 160	1.00	12	0.97
2000 320	0.50	8	0.96
4000 640	0.25	4	0.96

Definitions of τ_1 and ϕ_1 are the delay time and amplitude, respectively of the first peak of the ACF shown in the Figure 1.3 (Chapter 1). The pitch can be described by τ_1 of the ACF (Gelfand 1998). An example of determination of τ_e for the bandpass-noise stimuli is shown in Figure 4.4. The value of τ_e represents a kind of repetitive feature or reverberation contained within the source signal itself.

The measured and calculated values of τ_1 and τ_e as a function of the bandwidth and as a parameter of center frequency are shown in the Figures 4.5 and 4.6, respectively. Such measured values of τ_1, ϕ_1 and τ_e for the six center frequencies with their selected bandwidths are shown in the Table 4.1. The value of τ_1 corresponded to the center frequency of the stimuli. The value of τ_e increased as the

bandwidth decreased and was larger at lower center frequencies. The factors of ACF were calculated (APPENDIX - C).
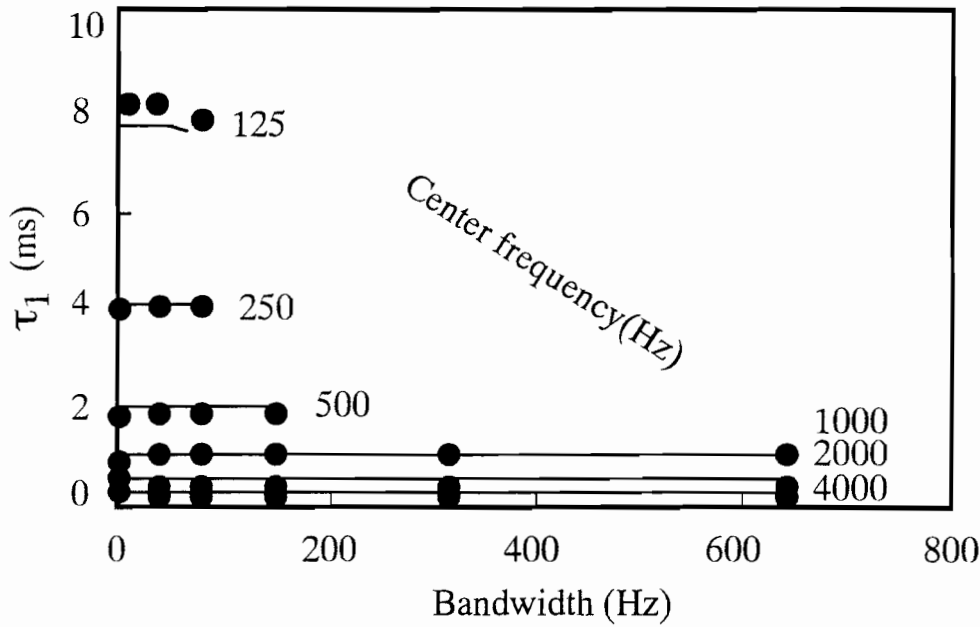
Figure 4.5. Measured (●) and calculated (—) values (see APPENDIX - C) of the bandpass noises as a function of bandwidth and as a parameter of center frequency.

4.2.3. Stimuli presentation

The duration sensations (DS) of bandpass noises were measured by the paired-comparison test. The subject was seated in the sound proof chamber on a chair in front of the loudspeaker. The distance of the center position of the head and the loudspeaker was fixed about 74 $\pm$1 cm. The first stimulus (duration of 150 ms) was a bandpass noise belonging to the combinations of six center frequencies (125, 250, 500, 1000,

2000 and 4000 Hz) and different bandwidths or a white noise in one session. The second stimulus was the white noise only. The duration of the second stimulus was randomly varied in the range of 140–230 ms in 10-ms steps.
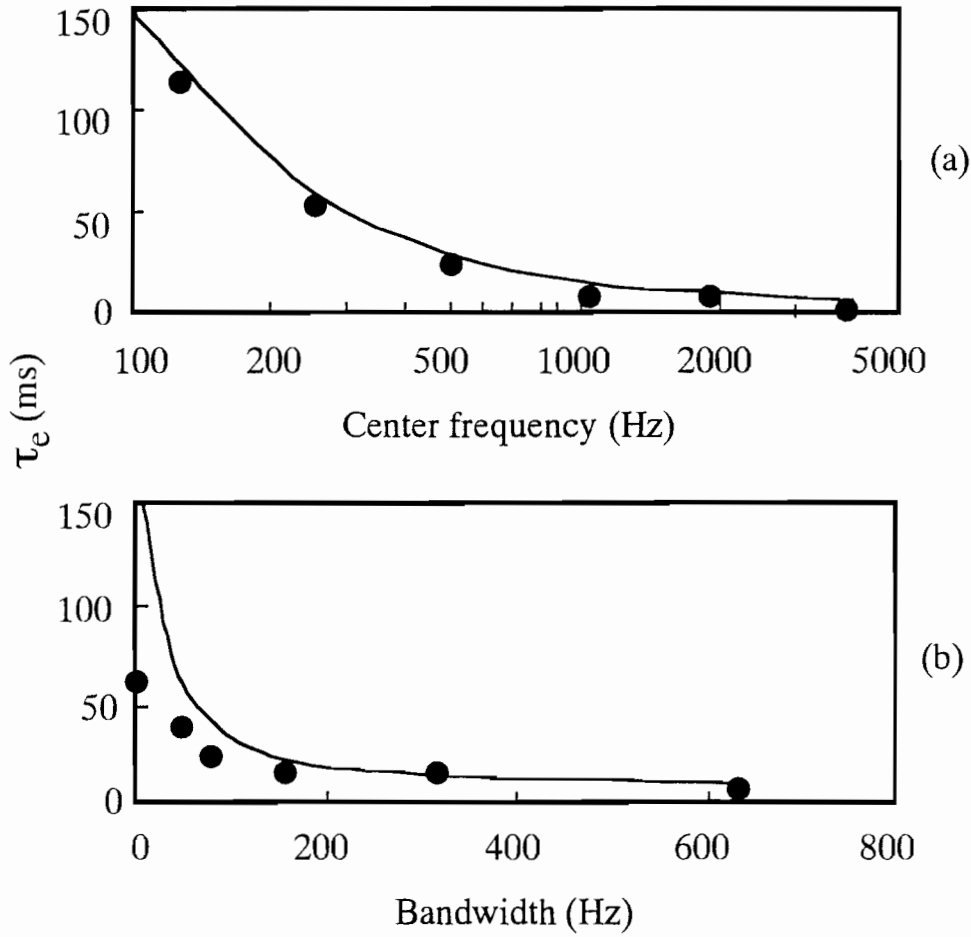

Figure 4.6. Measured (●) and calculated (—) values (see APPENDIX - C) of τ_e of the bandpass-noises as a function of bandwidth. (a) Six different center frequencies with selected bandwidths. (b) Six different bandwidths centered on 1000 Hz.

4.2.4. Subjective test

Ten subjects with normal hearing ability participated. Each subject was asked to judge whether the duration of the second of a pair of stimuli was longer or shorter than that of the first stimulus and to push the particular button only when the second stimulus seemed longer than first. The intra-pair and inter-pair gaps (for subjective judgments) were 1 and 3 s, respectively. Each pair was presented 20 times during each session.
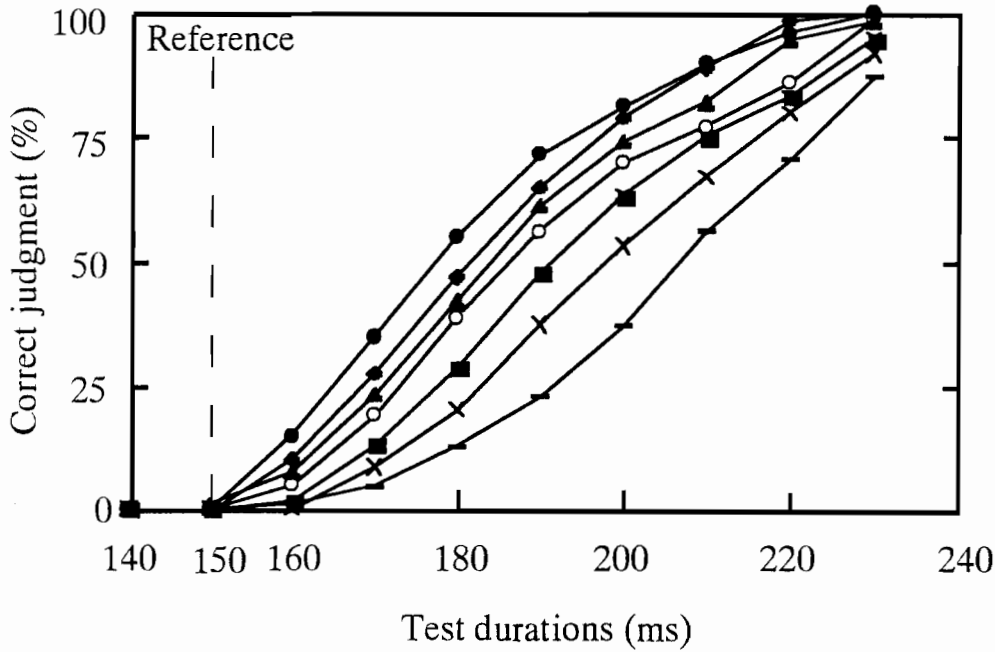
Figure 4.7. Cumulative frequencies of correct judgements of duration in the paired-comparison test for seven stimuli. The stimuli were six bandpass-noises with center frequencies. —: f_c = 125 Hz; $\times$: f_c = 250 Hz; ■: f_c = 500 Hz; ○: f_c = 1000 Hz; ▲: f_c = 2000 Hz; ◆: f_c = 4000 Hz and ●: white-noise. These bandpass noises have selected bandwidths (Δf) of 80, 80, 80, 160, 320 and 640 Hz, respectively.

4.3. Results

Twenty responses to each second stimulus-duration were obtained. The cumulative frequencies of the correct judgment of duration for the stimuli with the selected bandwidth of the six center frequencies and a white noise are shown in Figure 4.7.
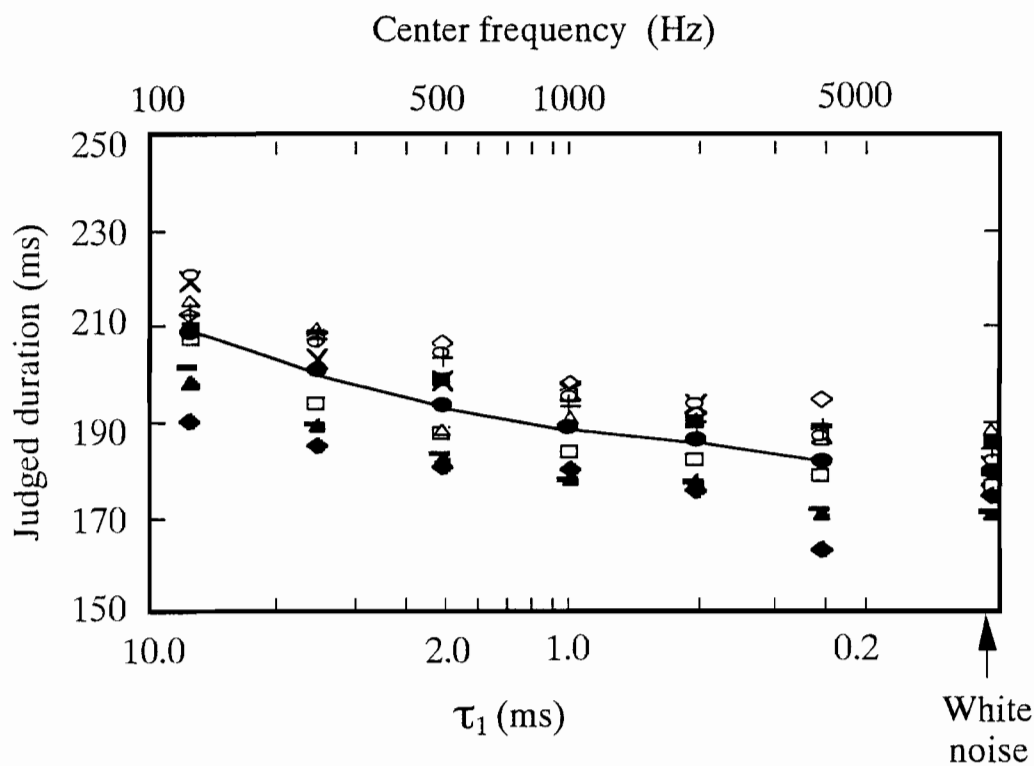
Figure 4.8. Judged duration for bandpass noises centered on the six different center frequencies with their selected bandwidths and white noise. The symbols indicate different subjects: ◇: HT; □: YK; △: KT; ×: KK; ■: TH; ○: KS; +: RS; —: SS; ▲: OY; ◆: NK; and ●: average.

The durations of the tests were found on the 50% line were 208.3, 199.7, 192.9,

188.4, 185.5, 181.1 and 178.5 ms. The percentages of the judged-longer duration than the standard duration (150 ms) are found respectively, 38.9 %, 33.1 %, 28.6 %, 25.6 %, 23.7 %, 20.7 %, and 19.0 %. The reference physical duration is 150 ms, and the 50% line of correct judgments is defined to be the duration sensation (DS). The results of ANOVA for judged-durations are listed in Table 4.2. The judged-durations for each individual and average for the durations of the six stimuli of bandpass noises were significantly longer (p $<$ 0.01) than that of the white-noise stimulus as shown in Figure 4.8 (standard deviation = 10.55 ms).

Table 4.2. F-values of the analysis of varience for the judged-durations among the stimuli of white noise and bandpass noise of six center frequencies (f_c) with the selected bandwidths (Δf).

f_c (Hz) = Δf (Hz) =	125 80	250 80	500 80	1000 160	2000 320	4000 640	WN
f_c (Hz) = 125 Δf (Hz) = 80	—	3.99	12.38**	24.76**	35.87**	38.77**	68.71**
250 80		—	2.41	7.97**	13.87**	18.09**	34.64**
500 80			—	1.21	3.60	7.34*	15.20**
1000 160				—	0.69	3.19	9.32**
2000 320					—	1.27	5.42*
4000 640						—	0.49
WN							—

* and ** indicate 5% and 1% significant level, respectively. WN: white noise.

4.4. Discussion

The longer duration sensations (DS) were found due to the lower frequencies (larger τ_1) of the bandpass-noise stimuli. This result is similar to those of the previous study on the duration sensation of pure-tone stimuli of differing frequency (Saifuddin et al. 2001a). Thus, the higher center frequency (smaller τ_1) produced shorter DS as the lower center frequency (larger τ_1) produced a longer DS (Figure 4.9). In the Figure 4.9, the relation between DS and τ_1 is well described by

$$\Delta DS \approx \alpha (\log \tau_1) + \beta, \qquad\qquad (4.1)$$

here, $\alpha \approx 15$ $\beta \approx 10$.
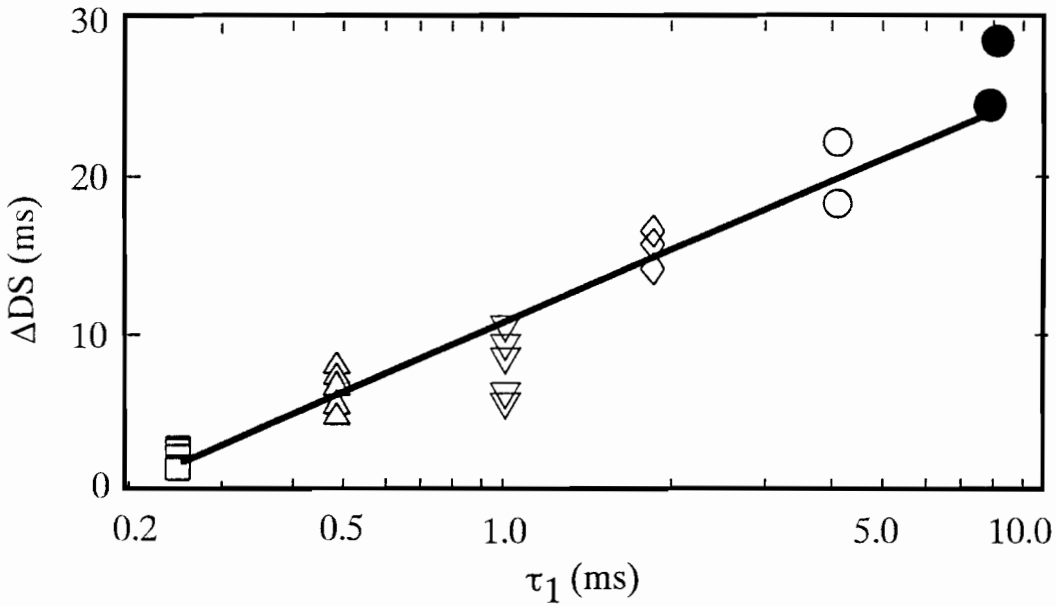
Figure 4.9. Judged greater-duration sensation (ΔDS) than the white noise as a function of measured τ_1 for all stimuli. Symbols indicate different center frequencies (f_c) or τ_1 : ● : 125; ○: 250; ◇: 500; ▽: 1000; △: 2000 and □: 4000 Hz with the combinations of different bandwidths.

In the Figure 4.9, percentile value ($\Delta_{DS}/150$) for example, at $\Delta_{DS} = 30$ is 20%. The DS is described by the value of τ_1, extracted from the ACF which, can be expressed by

$$DS = f(\tau_1). \qquad (4.2)$$

The judged greater durations for all of the bandpass-noise stimuli, in reference of the white noise as a function of $10\log\tau_e$ are shown in Figure 4.10. In the Figure 4.10 variation of the measured values of τ_e indicate variation of the DS within each center frequency. But, such τ_e also includes τ_1 in the Figure 4.10 as same as the Chapter 2 and 3. So it can be said that the τ_e has a minor effect on the DS.

$$DS \approx f(\tau_e), \qquad (4.3)$$

Here, f is the function and smaller τ_e indicates minor effect.
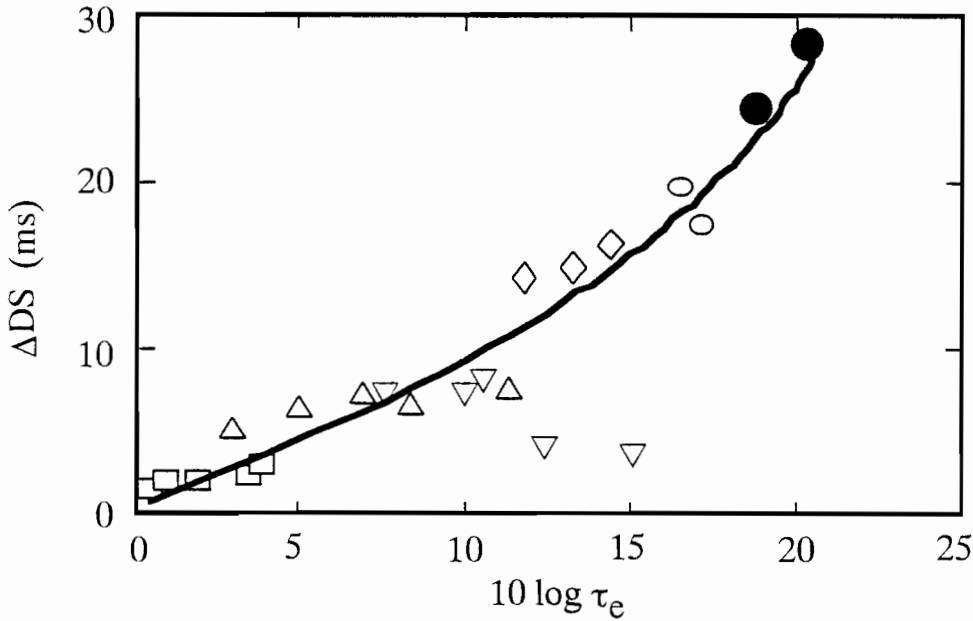
Figure 4.10. Judged greater-duration sensation (ΔDS) than the white noise as a function of $10 \log\tau_e$ for all stimuli. Symbols indicate different center frequencies: ●: 125; ○: 250; ◇: 500; ▽: 1000; △: 2000 and □: 4000 Hz with the combinations of different bandwidths.

In the Figure 4.11, plotted curve indicates responsibility of ϕ_1 for the variation DS which, can be expressed by

$$DS \approx f(\phi_1) \qquad\qquad (4.4)$$

Here, f is the function and smaller ϕ_1 indicates minor effect.
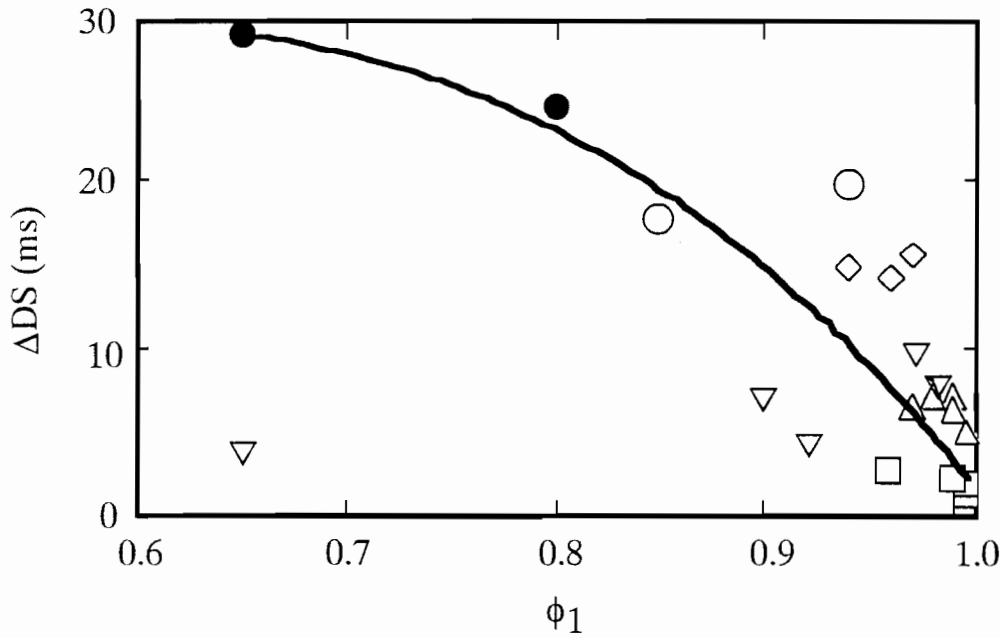
Figure 4.11. Judged greater-duration sensation (ΔDS) than the white noise as a function of measured ϕ_1 for all stimuli. Symbols indicate different center frequencies: ●: 125; ○: 250; ◇: 500; ▽: 1000; △: 2000 and □: 4000 Hz with the combination of different bandwidths.

The DS is also found here as the function (f) of the physical duration (PD) of the stimuli (Figure 4.7).

Results from the three figures (4.9, 4.10, 4.11), it can be explained that the three factors of ACF are the function of DS which, can be formulated by

$$DS = f(\tau_1, \tau_e, \phi_1, PD),\qquad\qquad (4.4)$$

here smaller symbols mean the minor effect on the DS.

4.5. Conclusions

The significant results are cited below.

1. The duration sensation (DS) of bandpass-noise stimuli used in this study (combination of six center frequencies and different bandwidths) were judged to be longer than that of a white-noise stimulus with the same sound pressure level (80 dB(A)) and duration.

2. The DS of bandpass-noise stimuli with lower center frequencies or lower pitch (larger τ_1) were judged to be longer than those stimuli with higher center frequency or higher pitch (smaller τ_1).

3. In the results, DS was found effected by the ϕ_1 as similar effects of ϕ_m (Chapter 3) where, ϕ_m includes ϕ_1.

4. Another minor effect of τ_e on the DS was found.

5. The effects of PD on the DS is also found here.

CHAPTER 5

DS IN RELATION TO THE FACTORS τ_1, ϕ_1, τ_e AND $\Phi(0)$ OF THE ACF OF BANDPASS-NOISE STIMULI UNDER REVERBERATION CONDITIONS

5.1. Preface

The purpose of the previous study (Ando et al. 2001b) is to examine the effects on duration sensation (DS) of the center frequency having various bandwidths of bandpass noise. Seven paired-comparison tests were conducted using white noise and six bandpass-noise stimuli with center frequencies of 125, 250, 500, 1000, 2000, and 4000 Hz, and their selected bandwidths (Zwicker et al. 1957) 80, 80, 80, 160, 320, and 640 Hz, respectively. The results of the tests indicated that the white-noise duration was judged to be of a shorter sensation than that of the bandpass-noise duration and a larger τ_1 (τ_1, being the time delay of the first peak of the autocorrelation function) was judged as a longer DS for the bandpass noises. In that study, stimuli were presented by a single loudspeaker in a soundproof chamber without reverberation effects and the subjects were seated alone at the horizontal distance of 74 ± 1 cm from the single loudspeaker. The sound pressure levels (SPL) as defined by the $\Phi(0)$ in the autocorrelation function (ACF) were kept constant at 80 dB(A). In order to that, another similar study was conducted here under a reverberation condition in a conventional

concert hall (Uhara Hall) to examine the reverberation effects and the effects of the SPL ($\Phi(0)$) on the DS.

A theory of the primary sensation for environmental noise -loudness, pitch, and timbre- was described by Ando (2001a). Such a primary sensation of environmental noise can be well described by the factors extracted from the ACF (Ando 1998). Saifuddin et al. (2001a) investigated the DS for stimuli consisting of white noise having a different amplitude (SPL) and a pure tone. Thus, the attenuation level of the envelope of the ACF of the stimuli (ϕ_m) was changed by the amplitude difference between the white-noise and pure-tone stimuli. The findings showed that the duration of stimuli with smaller ϕ_m was judged as being, longer sensation than that of larger ϕ_m. In the same study, the judgment of a longer-duration of sensation for the pure tone was found when the frequency was lower (longer τ_1) of the tone. To determine the DS as the primary one this study is designed while the reverberation situation represents environmental phenomena.

In this connection Oxenham reported in 2000 on a study of auditory-gap detection in which a larger deterioration was observed when two markers of broadband noise occupied different spectral regions but had the same fundamental frequency. According to the classical theory, the loudness of a bandpass noise remains constant as the bandwidth of the noise increases until the noise reaches the "critical band", after which the loudness increases with further increases of bandwidth (Zwicker et al. 1957, Zwicker et al. 1972, Greenwood 1961, and Hellman 1970). Merthayasa et al. (1994 and 1996) found that the loudness of sharply filtered noise (1080 dB/oct.) centered on 1000

Hz within the critical band is not constant. Florentine et al. (1996) compared the loudness of 1000 Hz tones with that of broadband noise over a wide range of levels. They found that the amount of temporal integration, defined as the level difference between equally loud stimuli with durations of 5 and 200 ms, varied with the level. Fujii et al. (2001) found that the noise of flying aircraft, which is a mixed state between 'tonal noise' and 'un-tonal noise', can be well represented by the factors extracted from the ACF and the interaural cross-correlation function. Such aircraft noise may be perceived to behave like a bandpass noise at a lower center frequency.

5.2. Methods

5.2.1. Physical situation of Uhara Hall

The well-known Uhara Hall (not soundproof) in Kobe, a conventional concert hall, was used in this study. The hall is finished in wood and contains 650 seats. The volume of the hall is 4870 m^3. One loudspeaker as a source was located at the center of the stage, 1.0 m above floor. The subjects were seated around the frontal side of the central position of the hall. The height of 1.1 m above the floor level corresponds to the ear positions. The distance between the source and the center ear positions of the subjects was 12.5 m. All of the physical conditions are shown in Figure 5.1 (a, b and c).
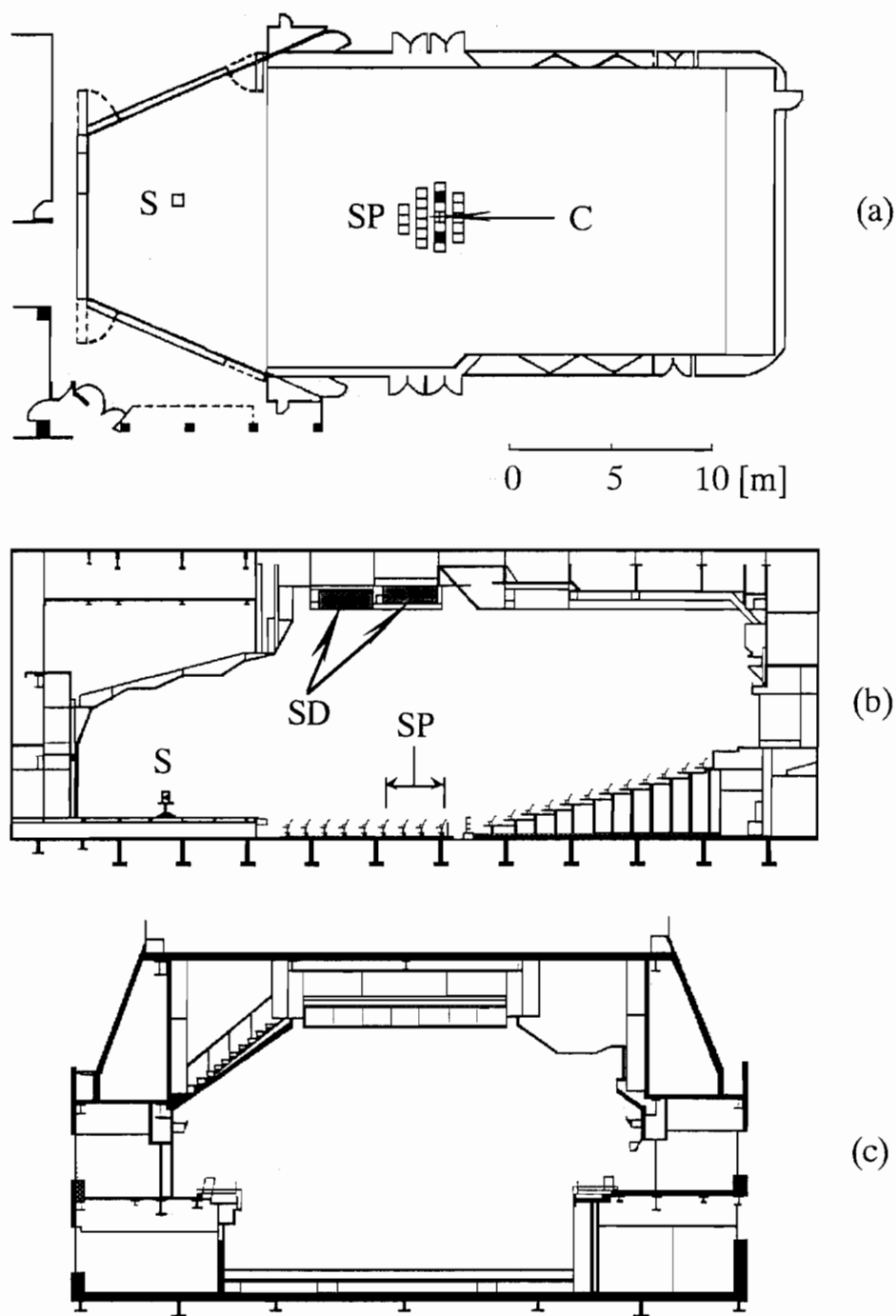
Figure 5.1. Seating positions 'SP' of the subjects are indicated in plan (a) and cross-section (b) and (c) of Uhara Hall, Kobe. Source position ' S' and center hearing position 'C' indicate where the SPLs were set up and the rise-fall times were measured. Impulse responses were measured at the center hearing position. SD: schroeder diffuser.

5.2.2. Stimuli preparing

Five bandpass-noise stimuli were simulated with a cut-off slope of 2068 dB/octave using two distal filters (48 and 140 dB/octave) through a series connection. The center frequencies 250, 500, 1000, 2000, and 4000 Hz were determined with the selected bandwidths 80, 80, 160, 320, and 640 Hz, respectively. The rise and fall times were defined by the time for a −3 dB drop from the steady level for all stimuli (Figure 5.2) and were measured at the center of the listening positions shown in the Table 5.1.
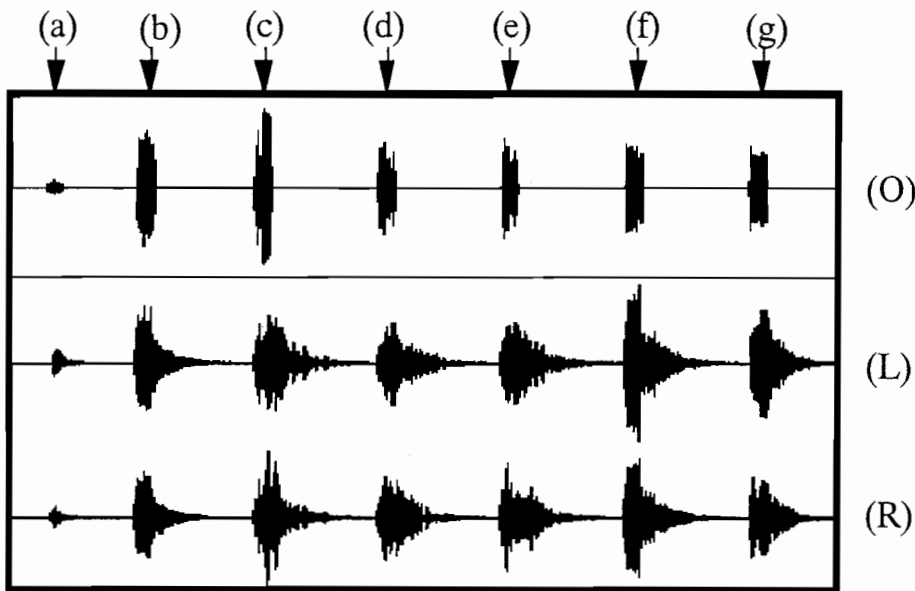
Figure 5.2. Examples of stimuli in which 'O' is the original signal without reverberation effects. The stimuli of 'L' and 'R' indicate the left-ear and right-ear hearing positions at 'C', respectively with reverberation. (a), (b), (c), (d), (e), (f) and (g) indicate WN 50 dB(A), WN 80 dB(A), BN (f_c: 250; Δf: 80), BN (f_c: 500; Δf: 80), BN (f_c: 1000; Δf: 160), BN (f_c: 2000; Δf: 320), and BN (f_c: 4000; Δf: 640). WN: white noise, BN: bandpass noise, f_c: center frequency and Δf: bandwidth. Each stimulus had a constant 150-ms duration.

5.2.3. ACF measurement

The SPLs ($\Phi(0)$ of the ACF) of all the direct sounds were measured as 80 dB(A) at the center of the subject's seating positions, except the SPL (50 dB(A)) of one white-noise stimuli presented first in the pair for the last (7 th) session. The factors τ_1, ϕ_1 and τ_e of ACF are defined in Figures 1.3 and 1.4 (Chapter 1), respectively, and were measured at the center listening position, as shown in Table 5.2. The theoretical procedure of the calculation of τ_1, ϕ_1 and τ_e is shown in the APPENDIX - C.
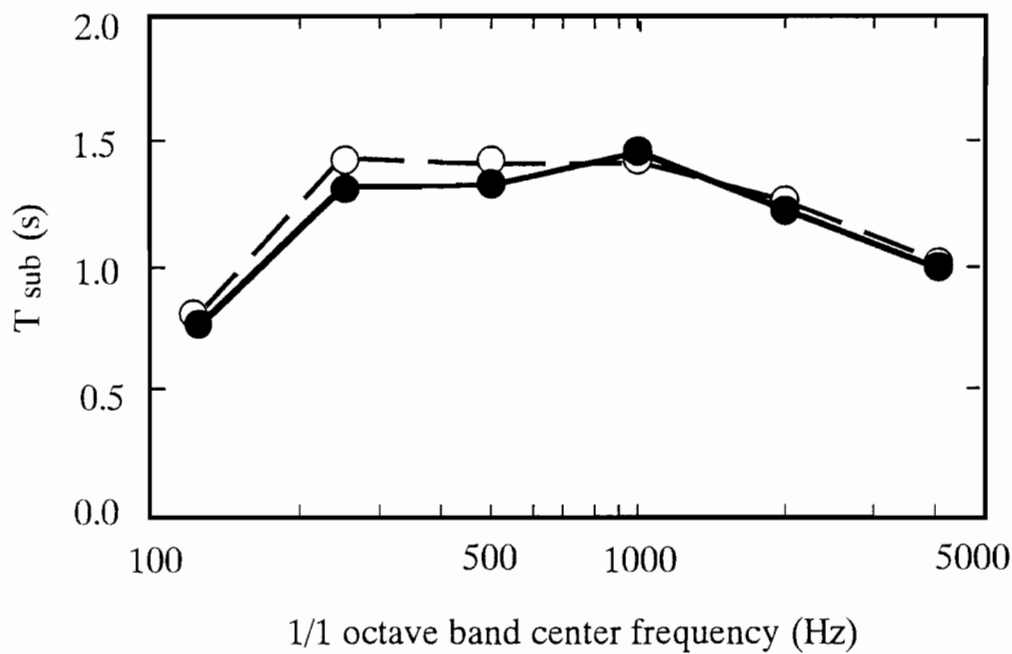
Figure 5.3. Values of reverberation time (Tsub) obtained by the binaural impulse responses at the center of the subject's seating area. '●' and '○' indicate the left and right hearing positions, respectively.

5.2.4. Impulse response measurement

Impulse responses were measured at the listening position within the subject's seating area (Figure 5.1). The measured reverberation times (Sabine 1900) are shown as the 'T_{sub}' in the Figure 5.3.

Table 5.1. The rise and fall times of the seven stimuli in which five bandpass and two white noises are shown under reverberation condition. The f_c: center frequency and Δf: bandwidth of the bandpass-noise stimuli. WN1 and WN2 indicate white noise with the SPL of 80 dB(A) and 50 dB(A), respectively.

Stimuli		White noise	Rise time (ms)	Fall time (ms)
f_c (Hz)	Δf (Hz)			
250	80		2.3	50.7
500	80		2.4	50.8
1000	160		2.5	51.3
2000	320		2.5	51.5
4000	640		2.6	52.2
		WN1	2.6	51.9
		WN2	2.3	50.8
Average			2.5±0.2	51.3±1.0
Standard deviation			0.14	0.59

5.2.5. Stimuli presentation

Seven stimuli consisting of two white noises and five bandpass noises were used in the paired-comparison tests conducted over seven sessions. Each session lasted 4 m and there was a 2 m rest period between sessions except for a 10 m rest period between the

4th and 5th sessions. The duration of the stimuli presented first in the pair was 150 ms. White-noise stimuli was presented ten times in durations of 140 to 230 ms at a step size of 10 ms at second in the pair. Each pair was presented five times randomly. The SPL of all the stimuli were kept constant at 80 dB(A), except for the white-noise stimulus (50 dB(A)) presented first in the last (7th) session. The intra-pair and inter-pair (response time) gaps were 1 and 3 s, respectively.

Table 5.2. Calculated (C) and measured values of τ_1 and ϕ_1 for the five bandpass noise stimuli with five center frequencies (f_c) and their selected bandwidths (Δf). W and R indicate without reverberation and reverberation conditions of measurement, respectively.

Stimuli		τ_1 (ms)			ϕ_1		
f_c (Hz)	Δf (Hz)	C	W	R	C	W	R
250	80	4.00	4.00	4.20	0.92	0.94	0.94
500	80	2.00	2.00	2.10	0.95	0.96	0.96
1000	160	1.00	1.00	1.00	0.97	0.97	0.97
2000	320	0.50	0.50	0.50	0.97	0.97	0.97
4000	640	0.25	0.25	0.25	0.97	0.97	0.97

5.2.6. Subjective test

Twenty-one subjects were asked to judge whether or not the duration of the second of a pair of stimuli seemed longer than the first stimulus. The subjects had to respond by a "tick mark" ($\sqrt{\ }$) on the answer sheet when the duration of the second stimulus seemed

longer than that of the first, otherwise they had to make a "cross-mark" ($\times$).
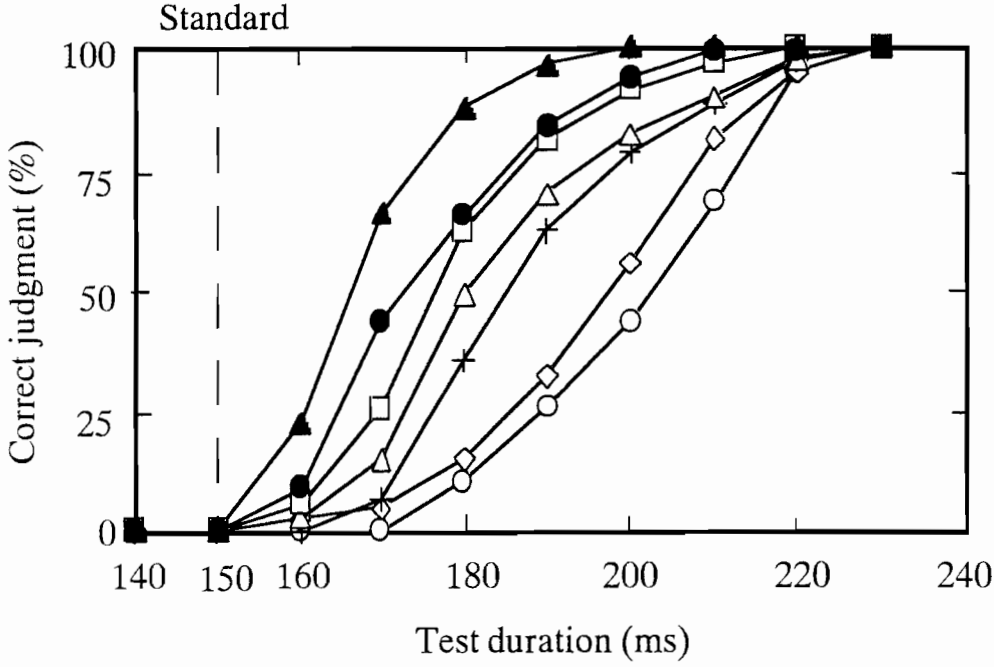
Figure 5.4. Cumulative frequencies of correct judgements of duration in the paired-comparison test for seven stimuli. The stimuli were five bandpass noises with the different center frequencies (f_c) and two white noises with two sound pressure levels. Symbols signify: $\bigcirc$: $f_c = 250$ Hz; $\Diamond$: $f_c = 500$ Hz; $+$: $f_c = 1000$ Hz; $\triangle$: $f_c = 2000$ Hz; $\square$: $f_c = 4000$ Hz; $\bullet$: white noise 80 dB(A), and $\blacktriangle$: white noise 50 dB(A). The bandpass noises have selected bandwidths (Δf) of 80, 80, 160, 320, and 640 Hz, respectively.

5.3. Results

The source stimulus was radiated from the loudspeaker on the stage and was measured as impulse responses by two small microphones placed at the two-ear entrances of a subject. The factor T_{sub} of the left hemisphere specialization (Ando et al. 1982) were measured and are shown in Figure 5.3.
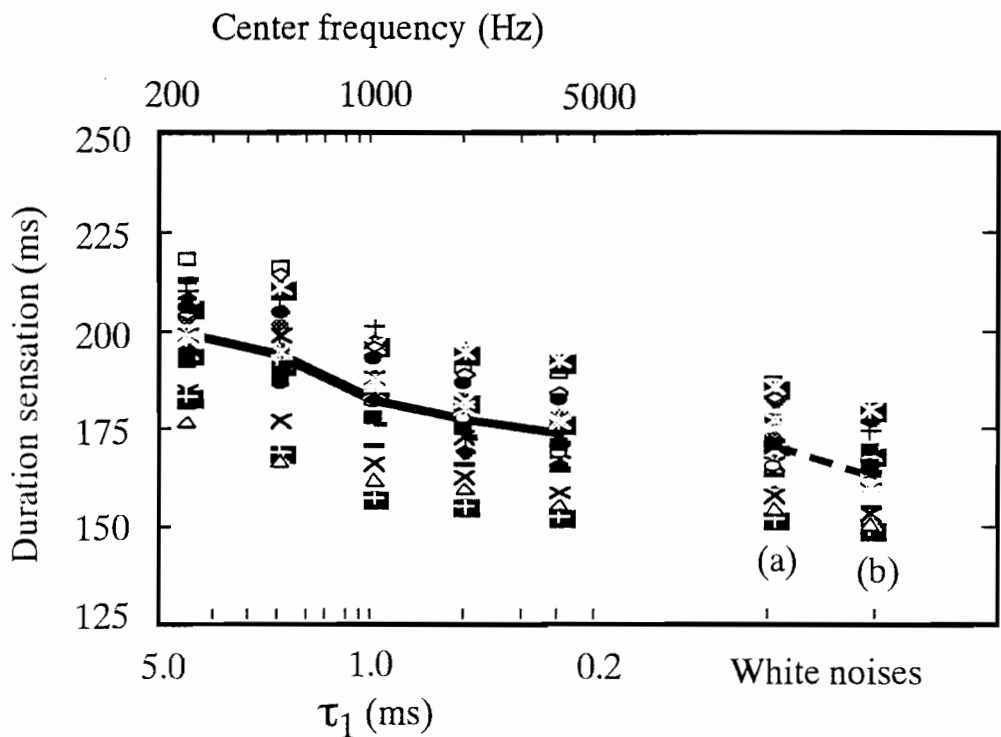
Figure 5.5. Judged duration for bandpass-noises centered on the five different center frequencies with their selected bandwidths and two white noises. The symbols indicate individual differences of duration sensation by twenty-one subjects (FK, HS, IS, JA, KF, KG, KK, KO, MT, NS, NY, OY, RS, SS, TH, TK, TM, TY, YK, YS, and RM). ▬: average.

The average value of the five responses at each pair for each subject out of 21 was obtained. The cumulative frequencies of the correct judgments of the subjects for the seven different stimuli presented first in the pair are shown in Figure 5.4. The reference physical duration was 150 ms, and the 50% line of correct judgments was defined to be the duration sensation (DS). The values on the 50% line of the cumulative frequencies of the averaged correct judgments are found to be 202.9, 197.5, 185.6, 180.6, and 176.9 ms for the bandpass-noise stimuli, whose center frequencies were 250, 500, 1000, 2000,

and 4000 Hz respectively, of their selected bandwidths. The percentages of the judged-longer duration than the standard duration (150 ms) are found respectively, 35.3 %, 31.7 %, 23.7 %, 20.4 %, and 17.95 %. The values for the white noises were found to be 173.0 and 166.0 ms respectively (longer; 15.3 and 10.6 %), with the SPL of 80 and 50 dB(A). The 50% line of cumulative frequencies of correct judgment of subjects with their individual (21 subjects) differences (standard deviation = 17.19) are shown in Figure 5.5.

Table 5.3. F-values of the analysis of varience for the judged durations among the stimuli of white noise (80 and 50 dB(A)) and bandpass noises of five center frequencies (f_c) with their selected bandwidths (Δf).

f_c (Hz) = 250 Δf (Hz) = 80	500 80	1000 160	2000 320	4000 640	WN 80 dB	WN 50 dB
f_c (Hz) = 250 Δf (Hz) = 80	1.77	20.11**	35.38**	49.15**	74.53**	133.11**
500 80	—	7.41**	15.64**	23.63**	37.01**	68.61**
1000 160		—	1.52	4.66*	10.89**	29.91**
2000 320			—	0.88	4.16*	17.60**
4000 640				—	1.11	10.03**
WN 80 dB					—	4.89*
WN 50 dB						—

* and ** indicate 5% and 1% significant level, respectively; WN: white noise.

5.4. Discussion

The duration sensations (DS) of bandpass-noise stimuli were found to be different due to the difference of the center frequency. It was clearly seen that the DS became gradually longer when the center frequencies belonging to a larger τ_1. The results of analysis of variance (ANOVA) for the judged-durations are listed in Table 5.3. The judged durations for each individual and the average for the durations of the five stimuli of bandpass noises were found to be significantly longer (p < 0.01) than those of the white-noise stimulus, as shown in Figure 5.5.
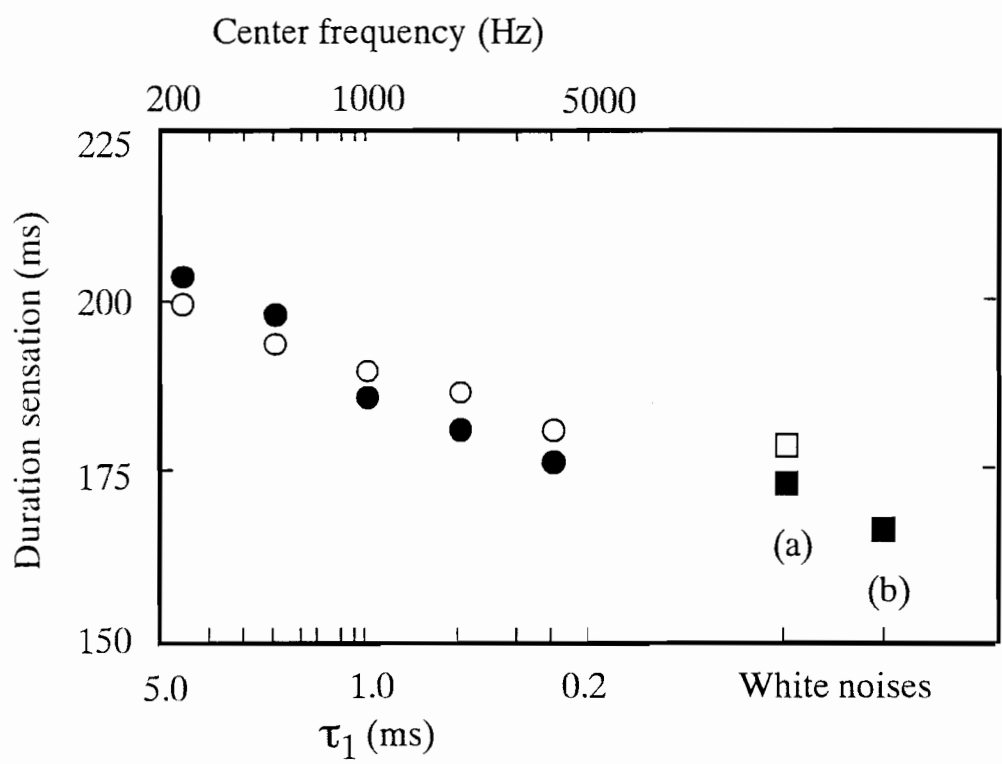
Figure 5.6. Duration sensation obtained by cumulative frequencies on the 50% line of correct judgments in the paired-comparison test for the five bandpass-noise and two white-noise stimuli. Black circle (●) indicates reverberation condition while, white circle (○) indicates without reverberation. (a): white noise 80 dB(A) and (b): white noise 50 dB(A).

This result was significantly similar (p $<$ 0.05) to those of the previous study of Ando et al. (2000b) on the DS for the same stimuli of bandpass noises and white noise without a reverberation condition (Chapter 4). In the Figure 5.6, the relation between DS and τ_1 is well described (for both studies) by

$$\Delta DS \approx \alpha (\log \tau_1) + \beta, \qquad (5.1)$$

here, $\alpha \approx 15$ $\beta \approx 10$.

The standard deviations of the individual responses for this study (21 subjects) was found to be 17.19 ms, whereas, for the previous study (10 subjects) it was 10.55 ms. Figure 5.6 shows that the DS of all the stimuli were found to be shorter under a reverberation condition, except for two stimuli with center frequencies of 250 and 500 Hz, than the DS without reverberation. It is important to understand that the DS is varied not by the effect of τ_e but by the effect of τ_1. This result was similar to that of previous study of Ando et al. (2000b). In Table 5.2, the values of ϕ_1 was found to have almost the same tendency as that of the previous study.

Results of another experiment show that the white-noise duration is judged as a longer sensation (p $<$ 0.05) when presented under higher SPL or higher $\Phi(0)$ (80 dB(A)) than that of the lower one (50 dB(A)), compared with the same white-noise durations (80 dB(A)). It makes it clear that the τ_1 and $\Phi(0)$ of the ACF are the approximate function of DS. That can be expressed by the below equation (5.1) in which, DS can be recognized as the fourth primary sensation of sound stimuli in the groups of loudness, pitch, and timbre.

$$DS = f[\tau_1, \tau_e, \phi_1, \Phi(0), PD], \qquad (5.2)$$

here, smaller symbol means minor effects on the DS.

5.5. Conclusions

The significant results of the study are cited briefly below.

1. The duration sensation (DS) of bandpass-noise stimuli (five center frequencies with selected bandwidths) were judged to be longer than that of a white-noise stimulus with the same sound-pressure level 80 dB(A) and duration under reverberation conditions.

2. The DS of the bandpass-noise stimuli with lower center frequencies or lower pitch (larger τ_1) was judged to be longer than those stimuli with higher center frequencies or higher pitches (smallter τ_1) under reverberation conditions.

3. The DS of a white-noise stimulus was found to be a longer sensation for the higher $\Phi(0)$ (SPL) than that of the lower one under reverberation conditions.

4. The DS of bandpass-noise stimuli of the five center frequencies 250, 500, 1000, 2000, and 4000 Hz with selected bandwidths of 80, 80, 160, 320 and 640 Hz, respectively were found to exhibit similar tendencies for both the reverberation and without reverberation conditions.

5. The PD is found as the function of DS.

CHAPTER 6

CONCLUDING REMARKS

6.1. Summary

Human sensations of the temporal duration were studied here. A series of experiments was conducted to examine the auditory sensations of the temporal duration using different sound stimuli. Most of the experiments were conducted in the sound proof chamber except one experiment in a conventional concert hall (chapter 5). Students as subject were taken part in these experiments. Psychological tests were performed to examine the auditory sensitivity of temporal durations. The paired-comparison method was used for the stimuli presentation and the response measurements. Subjects had to compare the durations of the two stimuli in each pair. The sound stimuli were selected on the basis of the measured factors of autocorrelation function (ACF) as parameter. The factors are τ_1, τ_e, ϕ_1, ϕ_m and $\Phi(0)$ explained in the Chapter 1. Sound stimuli used in the whole studies were white noise, pure tone (different frequencies), complex tone (different fundamental frequencies) and bandpass noise (different center frequencies and bandwidths). The responses were analyzed to examine whether or not the judged durations are effected by the factors of ACF. An "auditory-brain model" and a "theory of primary sensation" are used to interpret the data. The subjective judgments of the short-temporal durations are explained here as a

primary sensation and indicated as duration sensation (DS). All of the experiments are discussed in the Chapters 2 - 5. The significant findings of the results of the experiments are cited briefly below.

1. The DSs of pure-tone stimuli with larger τ_1 (lower frequencies) were found as longer than that of pure tone with smaller τ_1 (Chapter 2).

2. The DSs of the complex-tone stimuli with larger τ_1 (lower fundamental frequencies or lower pitch) were found to be longer than those stimuli with smaller τ_1 or higher fundamental frequency or higher pitches (Chapter 2).

3. The DSs were found almost same for the pure-tone and complex-tone stimuli with three stimulus-durations (140, 150, 160 ms) when, the τ_1 was 2 ms and found both of the DSs different from the pure tone with the τ_1 of 0.33 ms (Chapter 2).

4. The effects of τ_1 extracted from the ACF on the DS are found as the same tendency for the changing values of frequency, fundamental frequency, and center frequency, respectively, of the pure-tone, complex-tone, and bandpass-noise stimuli (Chapters 2, 4, 5).

5. The DSs of the mixed stimuli (pure tone and white noise) with higher values of attenuation level (larger ϕ_m) of the envelope of ACF were found as shorter than that of the smaller ϕ_m (Chapter 3).

6. The duration sensations (DS) of a white-noise stimulus were found to be longer than that of pure-tone stimuli (1000 Hz) under the same sound pressure level (80 dB(A)). That actually has the same duration (Chapter 3).

7. The DSs of bandpass-noise stimuli (combinations of six center frequencies and

different bandwidths) were found to be longer than that of a white-noise stimulus with the same sound pressure level and duration (Chapters 4, 5).

8. The DSs of bandpass-noise stimuli with lower center frequencies or lower pitch (larger τ_1) were found to be longer than those stimuli with higher center frequency or higher pitch (Chapter 4).

9. The DSs of bandpass-noise stimuli (five center frequencies of 250, 500, 1000, 2000 and 4000 Hz, with selected bandwidths) were found to be longer than that of a white-noise stimulus with the same sound-pressure level and duration under reverberation conditions (Chapter 5).

10. The DSs of the bandpass-noise stimuli with lower center frequencies or lower pitch (larger τ_1) were found to be longer than those stimuli with higher center frequencies or higher pitches under reverberation conditions (Chapter 5).

11. The DSs of a white-noise stimulus were found to be a longer sensation for the larger values of $\Phi(0)$ extracted from the ACF than that of the smaller one under reverberation conditions, in a room (Chapter 5).

12. The minor (non significant) effects of τ_e and ϕ_1 in the ACF on the DS were found (Chapters 4, 5)

13. The DSs of bandpass-noise stimuli of the five center frequencies 250, 500, 1000, 2000, and 4000 Hz with selected bandwidths of 80, 80, 160, 320 and 640 Hz, respectively, were found to exhibit similar tendencies for both the reverberation and without reverberation conditions (Chapters 4, 5).

14. Duration sensations (DS) of all the stimuli used in the study were found as

dependent on its physical duration (PD) (Chapters 2, 3, 4, 5).

Considering the whole experimental findings it can be concluded that "duration sensation (DS) of the sound stimuli is one of the primary sensations of the human-brain neuron which, is more or less responsible for the factors τ_1, τ_e, ϕ_1, ϕ_m, and $\Phi(0)$ extracted from the autocorrelation function". That can be expressed by

$$DS = f[\tau_1, \tau_e, \phi_1, \phi_m, \Phi(0), PD] \qquad (6.1)$$

Here, smaller τ_1, ϕ_1 and PD mean minor effects on the DS.

6.2. Applications

Information technology (IT) is become one of the most exiting output of science and technology. The new century hopefully will be engaged with the process of globalization and then 'IT' can play a vital-role on the key-point for the exchanges of information. In this reality, "high-tech" environments have been providing as a way, or even requiring us, to communicate with machines orally. In human-machine communications, machines are unable to intuitively understand our oral information. The machine has to learn the process of the human speech information. The oral exchanging technology of the man-machine system needs material and digital explanation of human speech. Autocorrelation function (ACF) can provide the effective factors, which are well concerned with the characteristics of speech signals.

The concluded studies were conducted using such of stimuli are recognized as the basic forms of sound signal (pure tone, complex tone, white noise, bandpass noise, etc.).

However, the sound signals frequently used in the verbal communications of the humans are supposed as the complex combinations of the basic forms. That is obviously speech-signal which is important for the information processing technology.

The knowledge of the duration sensation in relation to the factors extracted from the ACF can be used to develop technology of the speech signal processing. The basic components of the speech signals are sequenced on the physical time domain. Such physical duration can be varied for the auditory sensation in relation to the changing factors of the ACF in the signal. However, for the technology of digital conversion during out-put and in-put process of the signal such knowledge of the auditory duration sensation can be applied.

6.3. Further studies

On the basis of the present studies, wide possibilities of the future-research can be considered. One of the purposes of the future study is to find out the certain range of the values of particular factor of ACF, which is mostly concerned with the most preferred speech component. The four factors ($\tau_1, \tau_e, \phi_1, \Phi(0)$) extracted from the ACF can be measured for the speech signals. In this study, vowels, consonants and some standard syllables can be examined as basic components of speech signals. To measure the values of the factors with individual differences of the different vocal pronunciation, speech signals can be recorded onto a digital-audio tape (DAT). Then the signals will be reproduced through a D/A converter in the sound proof chamber. The microphone will

receive the signal and the then ACF can be analyzed. Most preferred vocal pronunciation can be identified by using preference-judgment test. To identify the most preferred vocal pronunciation of the components of the same speech signals; vowels, consonants and syllables can be used in the sound proof chamber in front of the subject. Last of all results from the measured ACF and psychological tests will be adjusted whether the individual physical-score is related to that of the most preferred speech stimuli.

Another important possibility of the future research is to find out the relationship between the certain-individual value of perceived duration and most preferred judgment of the particular speech signal. Considering all of the individual data whether or not the most preferred signal is related to that of the certain perceived duration can be examined. The factors of ACF of such speech signal can be measured in relation to that of the most preferred one.

It is also important to conduct study on the duration sensation between the sound duration and silent duration. The explanation of the sensation of silence duration could be a very significant knowledge when compared with the physical time duration. In this study the characteristics of markers of the particular silence duration can provide an effects on the duration sensation.

The duration sensations were examined on the direct comparison between pure tone and complex tone (chapter 2) with a fixed value of τ_1 (τ_1 = 2 ms). The contemporary study should be conducted using different values of τ_1 (e.g.; τ_1 = 4.0, 1.0, 0.5, 0.25 ms).

BIBLIOGRAPHY

Allan, L. G. (1979). The perception of time. Percept. *Psychophys*, **26**, 340-354.

Ando, Y. (1977). Effects of noise on duration experience. *Journal of Sound and Vibration*, **55**, 600-603.

Ando, Y. (1982). Perception of coloration in sound field in relation to the autocorrelation function. *J. Acoust. Soc. Am.*, **71**(3), 616-618.

Ando, Y., Otera, K. and Hamana, Y. (1983). Experiments on the universality of the most preferred reverberation time for sound fields in auditorium. *J. Acoust. Soc. Jpn.*, **39**, 89-95.

Ando, Y. (1985). *Concert Hall Acoustics*. Heidelberg: Springer.

Ando, Y. and Kang, S. H. (1987). A study on the differential effects of sound stimuli on performing left- and right-hemispheric tasks. *Acoustica*, **64**, 110-116.

Ando, Y., Yamamoto, K., Nagamatsu, H. and Kang, S. H. (1991). *Acoustics Letter*, **15**, 57-64.

Ando, Y. and Singh, P. K. (1996). A simple method of calculating individual subjective responses by paired-comparison tests. *Mem. Grad. School Sci. & Technol.*, Kobe Univ., **14-A**, 57-66.

Ando, Y. (1998). Architectural Acoustics, Blending Sound Sources, Sound Field, and Listeners. New York: *AIP Press/Springer*.

Ando, Y., Hiroko, T., Akio, M., Matsushima, T. and Saifuddin, K. (2000a). Subjective duration of every three-year periods for 3-18 years of age estimated by

students. *Journal of Human Ergology*, **28** (1).

Ando, Y., Saifuddin, K. and Sato, S. (2000b). Duration sensation when listening to bandpass noises. *J. Sound Vibration* (submitted).

Ando, Y. (2001a). A theory of primary sensations and spatial sensations measuring environmental noise. *J. Sound Vibration*, **241**(1), 3-18.

Ando, Y. (2001b). Orthogonal factors describing primary and spatial sensations of the sound field on a concert Hall. *International Conference on Acoustics*, Rome, 12 Sept. (to be presented).

Austin, J. and Marilyn, M. (1966). Perceived duration as a function of auditory stimulus frequency. *Journal of Experimental Psychology*, **71**, 358-364.

Creelman, C. D. (1962). Human discrimination of auditory duration. *J. Acoust. Soc. Am.*, **34**, 582-593.

Flanagan, J. L. and Saslow, M. G. (1958). Pitch Discrimination for Synthetic Vowels. *J. Acoust. Soc. Am.*, **30**, 435-442.

Florentine, M., Buus, S. and Poulsen, T. (1996). Temporal integration of loudness as a function of level. *J. Acoust. Soc. Am.*, **99**, 1633-1644.

Fraisse, P. (1961). L'influence de la frequence sur l'estimation du temps. *Anne'e Psychologica*, **61**, 325-339.

Fraisse, P. (1963). *The psychology of time*. New York: Harper and Row.

Fraisse, P. (1982). Rhythm and tempo. *In the Psychology of Music*. (Ed. Deutsch, D.). Ed., Academic Press, Orlando, Fl., Chapter 6.

Frankenhaeuser, M. (1959). *Estimation of time*. Stockholm: Almqvist & Wiksell.

Fujii, K., Soeta, Y. and Ando, Y. (2001). Acoustical properties of aircraft noise measured by temporal and spatial factors. *J. Sound Vibration,* **241**(1), 69-78.

Galambos, R. and Davis, H. (1943). The responses of single auditory nerve fibers to acoustic stimulation. *J. Neurophysiol,* **6**, 39-57.

Gelfand, S. A. (1998). *Hearing –An introduction to psychological and physiological acoustics.* Marcel Dekker, Inc, 353-368.

Greenwood, D. D. (1961). Auditory masking and the critical band. *J. Acoust. Soc. Am.,* **33**, 484-502.

Gyuau, M. (1902). *La genese de l'Idee de Temps.* Paris: Alcan.

Hall, J. L. (1983). A procedure for detecting variability of psychophysical threshold. *J. Acoust. Soc. Am .* **73**, 663-669.

Hellman, R. P. (1970). Effects of noise bandwidth on the loudness of 1000 Hz tone. *J. Acoust. Soc. Am.,* **48**, 500-504.

Hicks, R. E., Miller, G. W., Gaes, G. and Bierman, K. (1977). Concurrent processing demands and the experiencee of time-in-passing. Am. J. Psychol., **90**, 719-730.

Hind, J. E., Anderson, D. J., Brugge, J. E., and Rose, J. E. (1967). Coding of information pertaining to paired low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.,* **30**, 794-816.

Huggins, A. W. F. (1972). Just noticeable differences for segment duration in natural speech. *J. Acoust. Soc. Am.,* **51**, 1270-1278.

Inoue, M., Ando, Y. and Taguti, T. (2001). The frequency range applicable to pitch identification based upon the auto-correlation model. *J. Sound Vibration,* **241**, 105-

116.

Kato, H., Tsuzaki, M. and Sagisaka, Y. (1992). Acceptability and discrimination threshold for distortion of segmental duration in Japanese words. *Procedings of International Conference on Spoken Language Processing*, 507-510.

Kato, H., Tsuzaki, M. and Sagisaka, Y. (1993). Acceptability for durational modification of segments in words. Technical Report on Speech, *Acoust. Soc. Jpn.*, SP92-145.

Katsuki, Y., Sumi, T., Uchiyama, H. and Watanabe, T. (1958). Electric responses of auditory neurons in cat to sound stimulation, *J. Neurophysiol.*, **21**, 569-588.

Kimura, D. (1973). The asymmetry of the human brain. *Sci. Amer.*, **228**, 70-78.

Kiang, N. Y. S. (1965). *Discharge Pattern of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, MA).

Liberman, M. C. (1982). The cochlear frequency map for the cat: Labeling auditory nerve fibers of known characteristic frequency. *J. Acoust. Soc. Am.*, **72**, 1441-1449.

Martin, J. K., Kox, A. J., Renn, J., and Robert, S. (edt.) (1995). The collected papers of Albert Einstein. *Prinston University Press*, Prinston, 4.

Merthayasa, I. G. N., Hemmi, H. and Ando, Y. (1994). Loudness of a 1kHz pure tone and sharply (1080 dB/Oct.) filtered noise centered on its frequency. *Memoirs of Graduate School of Science and Technology, Kobe University*, **A 12**, 147-156.

Merthayasa, I. G. N. and Ando, Y. (1996). Variation in the autocorrelation function of narrow band noises: Their effect on loudness judgement. *Japan and Sweden Symposium on Medical Effects of Noise*, **3**.

Michon, J. A. (1965). Studies on subjective duration II. Subjective time measurement during tasks with different information content. *Acta Psychologica*, **24**, 205-219.

Newton ,I. (1687). *Philosophiae naturalis principia mathematica*, London.

Oxenham, A. J. (2000). Influence of spatial and temporal coding on auditory gap detection. *J. Acoust. Soc. Am.*, **107**, 2215-2223.

Pastor, M. A. and Artieda, J. (1996). Neurophysiological mechanism of temporal perception. *Time, Internal Clocks and Movements*. Amsterdam: Elsevier, 9-13.

Plomp, R. (1970). Timbre as a multidimensional attributes of complex tones. In R. Plomp and G. F. Smoorenburg (eds.), *Frequency Analysis and Periodicity Detection in Hearing*. Leiden: Sijitthoff, 397-414.

Robinson, D. E. and Watson, C. S. (1973). Psychophysical methods in modern psychoacoustics. In J. V. Tobias (ed.). *Fundamentals of Modern Auditory Theory*, **2**. New York: Academic Press, 99-131.

Rose, J. E., Brugge, J. F., Anderson, D. J. and Hind, J. E. (1967). Phase-locked response to low frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.*, **34**, 685-699.

Sabine, W. C. (1900). Reverberation. The American Architect and the Engineering Record (Sabine, W. C. Prefaced by Beranek, L. L. (1992). *Collected Papers on Acoustics*. Peninsula, Los Altos, CA, Chap. 1.

Saifuddin, K., Sakai, H. and Ando, Y. (2001a). Duration sensation of simulated environmental noise in relation to its autocorrelation function. *J. Sound Vibration*, **241**(1), 117-127.

Saifuddin, K. and Ando, Y. (2001b). Duration sensation when listening to bandpass noises in a room. *Journal of Temporal Design in Architecture and the Environment,* (reviewed).

Sato, S., Kitamura, T., Sakai, H. and Ando, Y. (2001). The loudness and pitch of "complex noise" in relation to the factors extracted from the auto-correlation function. *J. Sound Vibration,* **241,** 97-103.

Schouten, J. F. (1940). The Residue, a New Concept in Subjective Sound Analysis, *Proc. Kon. Ned. Akad.,* **43,** 356-365.

Schouten, J. F., Ritsma, R. and Cardozo, B. (1962). Pitch of the Residue. *J. Acoust. Soc. Amer.,* **34,** 1418-1424.

Seebeck, A. (1841). Beohachtungen uber einige Bedingungen der Entstehung von Tonen. *Ann. Phys. Chem.,* **53,** 417-436.

Sharon, M. Abel (1971). Discrimination of temporal gaps. *J. Acoust. Soc. Am.,* **52,** 519-524.

Sperry, R. W. (1974). Lateral specialization in the surgically separated hemispheres. *The Neurosciences*: Third study program (Eds. Schmitt, F.O., and Worden, F. C.). MIT Press, Camgridge, MA, Chapter 1.

Stanley, G. A. (1998). Psychoacoustic methods. *Hearing- an introduction to psychological and physiological acoustics.* New York: Marcel Dekker Inc., 246-249.

Stevens, S. S. and Volkman, J. (1940). The relation of pitch to frequency. *Am. J. Psych.,* **53,** 329-353.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychol. Rev.*, **34**, 273-289.

Vroon, P. A. (1970). Effects of presented and processed information on duration experience. *Acta Psychologica*, **34**, 115-121.

Wightman, F. L. (1973). Pitch and stimulus fine structure. *J. Acoust. Soc. Am.*, **54**, 397-406.

Woodrow, H. (1951). Time Perception. *Hand book of experimental psychology* (S. S. Stevens, editor), New York: Wiley.

Zakay, D., Nitzan, D. and Glicksohn, J. (1983). The influence of task difficulty and external tempo on subjective time estimation. *Perception and Psychophysics*, **34**, 451-456.

Zwicker, E., Flottorp, G. and Stevens, S. S. (1957). Critical Bandwidth in Loudness Summation. *J. Acoust. Soc. Am.*, **29**(5) 548-557.

Zwicker, E. and Fastl, H. (1972). On the development of the critical band. *J. Acoust. Soc. Am.*, **52**, 699-702.

APPENDIX - A

Ando and Singh (1996) developed a simple method of calculating scale values of individual preference using a single observation for a set of stimuli. The method is based on the law of comparative judgment of Thurstone (1927). It used the linear range of cumulative normal distribution between the probability and the scale values (Z-value) is shown in the Figure A.a. The method is used in the study (Chapter 2) to calculate the scale values of the differential duration judgments of the sound stimuli.
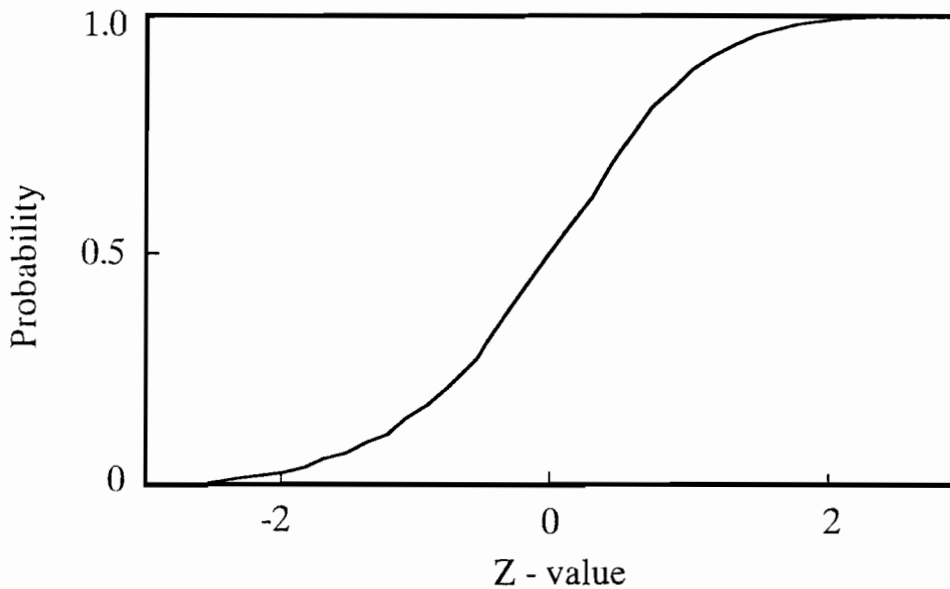
Figure A.a. The cumulative distribution (normal ogive) between the probability and scale value (Z - value). σ : standard deviation.

To calculate scale values of the comparative judgments of sound durations the method was used as an approximation for Case-V of Thurstone's law. Using linear

range $(0.05 \leq P \leq 0.95, \ P$: probability) of normal ogive according to the Figure A.a. The scale values of subjective judgments S_i is obtained from the equation below.

$$S_i \approx \frac{\sqrt{2\pi}(2T_i - N)}{2N} \qquad\qquad (A.1)$$

where T_i is the total number of scores of the expected judgments for the stimulus and N is the number of stimuli used in the paired-comparison tests.

A 'test of goodness of fit' also used in this method to justify the consistency of the scale values with the individual sound fields. The test considers a quantity λ as a ratio specified by

$$\lambda = \frac{\sum\limits_{(i,j)} |S_i - S_j| p}{\sum\limits_{(i,j)} |S_i - S_j|}, \qquad 0 \leq \lambda \leq 1 \qquad\qquad (A.2)$$

Here S_i and S_j are the subsequent scale values; and p means poorness of the consistency. So, considering an ordered pair such as (i, j), a solution for scale values should be consistent under the formulation if

$$S_i \geq S_j \qquad\qquad (A.3)$$

If $S_i > S_j$ means no error attributable to the scale values. Whereas if

$$S_i < S_j \qquad\qquad (A.4)$$

then $(S_j - S_i)$ denotes the error. So that conditions (A.3) and (A.4) are the measures of the goodness and poorness (p), respectively. The value of λ corresponds to the average error of the scale value which, should be smaller as less than 10%. If the value of λ is equal to zero, then the results are fitted consistantly.

The number of violations (poorness) is the k according to the condition expressed

by the above equations. The percentage of violations d is defined by

$$d = \frac{2k}{N(N-1)} \times 100, \qquad\qquad (A.5)$$

Thus, all values of d or k and λ can be considered as the test of goodness of fit in this method.

APPENDIX - B

The procedure for calculation of the peak amplitudes of the ACF envelope ϕ_m (m = 1, 2, 3, . . .) shown in Figure 3.1 is described here. The ACF of the sound stimuli $\Phi_p(\tau)$ can be defined as

$$\Phi_p(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} p(t)p(t+\tau)dt. \tag{B.1}$$

where τ is the delay time.

In this study, the sound stimulus was $p(t) = s(t) + An(t)$; where $s(t)$ is the pure tone, $n(t)$ is the white noise being incoherent with $s(t)$ and coefficient A is the relative amplitude of $n(t)$ to $s(t)$. Here, sound energies of $s(t)$ and $n(t)$ should be equal so as to have a relation: $\Phi_s(0) = \Phi_n(0)$. The integration interval 2T was 2.0 s. For such a mixed stimulus, normalized amplitude of m-th peak ϕ_m is given in the form

$$\Phi(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} [s(t)s(t+\tau) + a^2 n(t)n(t+\tau)]t.$$

$$= \Phi_s(\tau) + a^2 \Phi_n(\tau). \tag{B.2}$$

The normalized ACF $\phi(\tau)$ becomes

$$\phi(\tau) = \frac{\Phi(\tau)}{\Phi(0)} = \frac{\Phi_s(\tau) + A^2\Phi_n(\tau)}{\Phi_s(0) + A^2\Phi_n(0)}$$
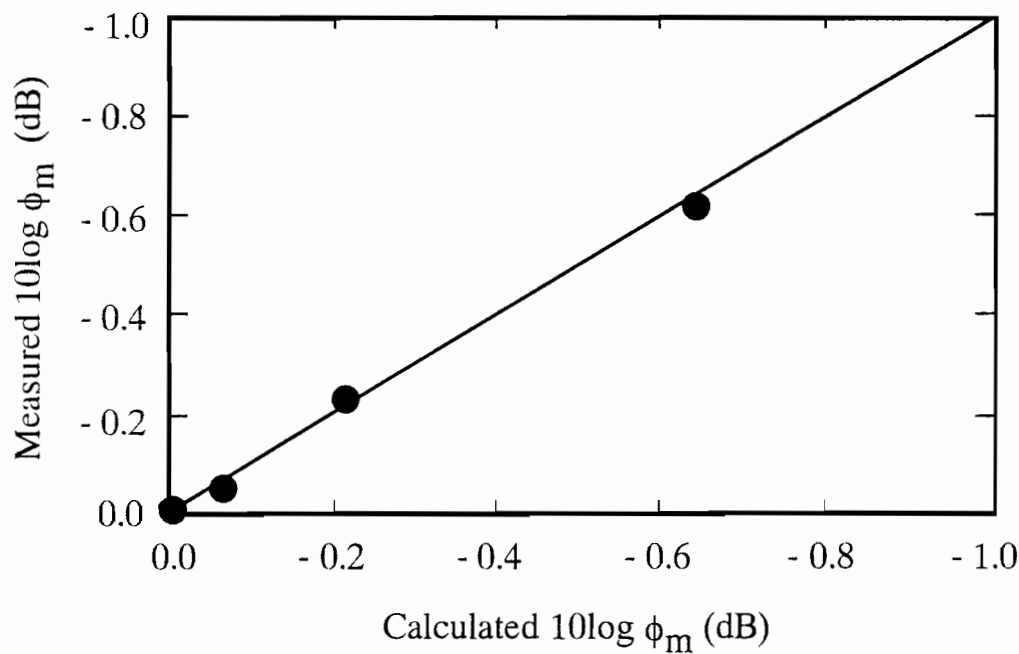
$$= \frac{\phi_s(\tau) + A^2\phi_n(\tau)}{1 + A^2},$$

where we put $\phi_s(\tau) = \frac{\Phi_s(\tau)}{\Phi_s(0)}$ and $\phi_n(\tau) = \frac{\Phi_n(\tau)}{\Phi_n(0)}$.

When $s(t) = \sin \omega t$ and $n(t)$ is the white noise,

$$\phi(\tau) = \frac{\cos \omega\tau + A^2\delta(\tau)}{1 + A^2}, \tag{B.3}$$

where $\delta(\tau)$ is the Dirac delta function. Thus, $\phi_{(0)} = 1$.

$$\phi_m = 10 \log_{10} \frac{1}{1+A^2}, \qquad m = 1, 2, 3, \ldots \qquad \text{(B.4)}$$

Figure B.a. The horizontal and vertical axes are, respectively, the calculated and measured ϕ_m for the envelope of the autocorrelation function for four stimuli.

Measured and calculated values of ϕ_m for four of the stimuli used in this study are listed in Table 3.1 and shown Figure B.a.

APPENDIX - C

The autocorrelation function (ACF) of the bandpass noises after passing through an ideal filter with a flat response between upper and lower frequencies f_2 and f_1 is given by (Ando, 1982)

$$\phi(\tau) = \frac{2}{\Delta\omega\tau}\sin(\frac{\Delta\omega\tau}{2})\cos(\frac{\Delta\omega_c\tau}{2}) \qquad (C.1)$$

where $\Delta\omega = 2\pi(f_2 - f_1)$, and $\Delta\omega_c = 2\pi(f_2 + f_1)$.

The envelope of the ACF of the bandpass noises is

$$2/\Delta\omega\tau \sin(\Delta\omega\tau/2), \text{ for } 0 \leq \Delta\omega\tau \leq \pi$$

$$\text{and } 2/\Delta\omega\tau, \text{ for } \Delta\omega\tau > \pi. \qquad (C.2)$$

The calculated values of τ_1 and τ_e as a function of the bandwidth and as a parameter of the center frequency are shown in Figures 4.5 and 4.6, respectively. Agreements between them are satisfactory, except for a few cases. Thus, we use measured values hereafter.

LIST OF PUBLICATIONS

The publications below were prepared on the basis of the Doctoral research. The brackets at the end of reference indicate the chapter's number where the paper is described in detail.

Full papers:

[1] Yoichi Ando, Hiroko Tsuruta, Akio Motokawa, Takaji Matsushima and **Kazi Saifuddin** (2000). Subjective Duration of Every Three-Year Periods for 3-18 Years of Age Estimated by Students. *Journal of Human Ergology*, 28.

[see Chapter 1]

[2] **Kazi Saifuddin**, Hirouki Sakai and Yoichi Ando (2001). Duration Sensation of Simulated Environmental Noise in Relation to its Autocorrelation Function. *Journal of Sound and Vibration*, 241, 117-127.

[see Chapters 2 and 3]

[3] **Kazi Saifuddin**, Takaji Matsushima and Yoichi Ando (2001). Duration Sensation when Listening to Complex Tone and Pure Tone. *Journal of Temporal Design -in architecture and the environment* (to be submitted).

[see Chapter 2]

[4] Yoichi Ando, **Kazi Saifuddin** and Shin-Ichi Sato (2001). Duration Sensation when Listening to Bandpass Noises. *Journal of Sound and Vibration* (reviewed).

[see Chapter 4]

[5] **Kazi Saifuddin** and Yoichi Ando (2001). Duration Sensation when Listening to Bandpass Noises in a Room. *Journal of Temporal Design -in architecture and the environment* (reviewed).

[see Chapter 5]

Proceedings of international congress:

[1] **Kazi Saifuddin**, Hirouki Sakai and Yoichi Ando (1999). Duration Experience During Listening to Simulated Environmental Noise in Relation to its Effective Duration of the Autocorrelation Function. *"1st International Workshop on New Systems for Identification and Evaluation of Regional Environmental Noise"* **(Kobe, Japan)**, 5.

[see Chapters 2 and 3]

[2] Yoichi Ando and **Kazi Saifuddin** (2000). Duration Experience During Listening to Bandpass Noises.*"4th Japanese-Swedish Noise Symposium on Medical Effects"* **(Stockholm, Sweden)**.

[see Chapter 4]

[3] **Kazi Saifuddin**, Takaji Matsushima and Yoichi Ando (2001). Duration Sensation When Listening to Complex Tones *"141st International Conference of Acoustical Society of America"* **(Chicago, USA)**, 4pPP6.

[see Chapter 2]