Paper:

Accuracy of Synchrony Judgment and its Relation to the Auditory Brainstem Response: the Difference Between Pianists and Non-Pianists

Eriko Aiba*1, Koji Kazai*1, Takayuki Shimotomai*2, Toshie Matsui*3, Minoru Tsuzaki*4, and Noriko Nagata*1

*1Kwansei Gakuin University
2-1 Gakuen, Sanda, Hyogo 669-1337, Japan
E-mail: aiba@kwansei.ac.jp
*2Brain Science Institute, Tamagawa University, Japan
*3Nara Medical University, Japan
*4Kyoto City University of Arts, Japan
[Received February 28, 2011; accepted July 26, 2011]

Synchrony judgment is one of the most important abilities for musicians. Only a few milliseconds of onset asynchrony result in a significant difference in musical expression. Using behavioural responses and Auditory Brainstem Responses (ABR), this study investigates whether synchrony judgment accuracy improves with training and, if so, whether physiological responses are also changed through training. Psychoacoustic experiments showed that accuracy of synchrony judgment of pianists was higher than that of non-pianists, implying that pianists' ability to perceive tones increased through training. ABR measurements also showed differences between pianists and non-pianists. However, cochlear delay, an asymmetric aspect of temporal processing in the human auditory system, did not change with training. It is possible that training improved ability related to temporal tone perception and that training may increase synchrony in auditory nerve firing.

Keywords: Auditory Brainstem Response (ABR), cochlear delay, pulse, pianist, synchrony

1. Introduction

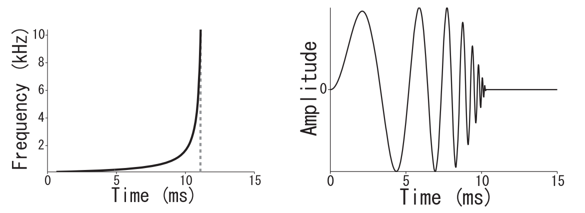
Onset synchrony is widely assumed to be an important cue for the perceptual unification of multiple sinusoidal sound components as a single tone [1]. The collapse of onset synchrony (onset asynchrony of sound components) results in various influences on the perception of the sound timbre, pitch and so on [2–4].

In the music performance field, onset asynchrony is considered an important factor in ensemble performance [5] and can change the expressive qualities of music [6]. Rasch [5] reported a mean value of asynchrony between tones of about 36 ms when musicians play in an ensemble, even when tones were nominally simultaneous, as in a chord. Such onset asynchronies enable both the performers and the listeners to hear the simultaneous tones as distinct from each instrument and are not considered performers' mistakes [5]. When two complex tones are played together, detection of higher tones in the presence of lower ones is strongly affected by onset asynchrony. For example, when lower tones are played 10 ms after higher tones, the higher tones are heard with up to a 10-dB reduction in detection threshold. When higher tones are played 30 ms before lower tones, the threshold for perception of higher tones is unchanged from when they are presented alone [6]. Our preliminary interview with pianists indicated that they try to control timing of tone onsets as if they were aware of the scientific findings of Rasch [5, 6]. They insisted that they use a technique in which they play certain tones slightly earlier than others, timing them to make them stand out (e.g., within a chord). As the tones must be heard as a chord, however, the degree of asynchrony should stay within the limits of "perceptual synchrony." Additionally, there is a similar performance technique called "arpeggio." An arpeggio is a group of tones that make up a chord played one after the other, consecutively and quickly, either rising or falling [7]. All the tones are consciously performed to express the asynchrony.

It is therefore important for musicians to control the extent of onset asynchrony. Just a few milliseconds of onset asynchrony will result in a significant difference in musical expression. Therefore, musicians are always carefully monitoring whether or not the onset of tones is simultaneous.

1.1. Cochlear Delay

Even if all of the components of a single complex tone physically begin simultaneously, their temporal relation is not preserved at the cochlear level because of its physical characteristics. The stiffness of the cochlear Basilar Membrane (BM) gradually decreases from the basal side (closer to the oval window) to the apical side [8]. There-



(a) function (b) waveform

Fig. 1. The solid line of the left panel (a) shows the instantaneously increasing frequency pattern as a function of time used under the compensated delay condition. The broken line shows the time required for all frequencies to reach maximum amplitude on the BM. The right panel (b) shows the waveform of the chirp with compensated delay that was calculated based on this function.

fore, the higher components of an input wave excite the basal side, while the lower components excite the apical side. Due to the time delay required for wave transmission, the vibration caused by low-frequency components reaches its associated location later than that caused by high-frequency components. This phenomenon is referred to as "cochlear delay." This delay largely occurs for components below 1000 Hz, and the vibration in the lowest-frequency associated location is delayed by about 10 ms relative to the vibration in the high-frequency associated location [9].

In our previous psychoacoustic study [10, 11], an experiment was performed to measure the accuracy of synchrony judgment using stimuli that controlled cochlear delay. The results of the experiment showed that the synchrony judgment accuracy was highest for stimuli that evoke an intrinsic cochlear delay. Furthermore, accuracy was higher for stimuli that evoke enhanced cochlear delay than for stimuli that cancelled out cochlear delay, which implies that there is an asymmetric aspect of temporal processing in the human auditory system. In the present study, a psychophysical experiment was designed to investigate whether musicians exhibit similarly asymmetric processing and whether musicians have more accurate synchrony judgment.

In this study, two types of chirps and a pulse were used as experimental stimuli to investigate whether or not cochlear delay imposes a systematic bias on the judgment of perceptual synchrony of two sounds. These are (a) a compensated delay chirp, (b) an enhanced delay chirp, and (c) a pulse that evokes an intrinsic cochlear delay. The compensated delay chirp instantaneously increased its frequency to cancel out the cochlear delay (**Fig. 1**). The three types of sound have flat continuous amplitude spectra but have different phase spectra. The two chirps can be technically categorized as a time stretched pulse. The word "pulse" in this paper is used in its strict definition, that is, a signal with a flat amplitude spectrum being added in a cosine phase.

The increasing frequency pattern function (**Fig. 1**), originally calculated by Dau et al. [12], was based on the one-dimensional, linear cochlear model of de Boer [13]. The stiffness of an object is largely responsible for the propagation speed of a travelling wave. Therefore, the basic assumption of this increasing frequency pattern is that the physical stiffness of the human BM decreases exponentially along the cochlear partition from base to apex. According to de Boer [13], the function of stiffness of the BM, $c(\boldsymbol{x})$, is as follows:

$$c(\boldsymbol{x}) = C_0 \cdot \exp(-\alpha \boldsymbol{x}) \quad \quad (1)$$

Here, $\boldsymbol{x}$ is a single function of coordination that represents the mass and damping of the cochlear partition. C_0 represents a value of stiffness at the stiffest place on the base of the BM. This value was chosen as 10^9 dyn/cm^3. The parameter α determines the rate of instantaneous frequency changes and was set at 3 cm^{-1}. These values were also derived by de Boer [13]. The compensated delay chirp had an increasing frequency pattern to "cancel out" the cochlear delay. Accordingly, it is assumed that the excitation along the BM arises simultaneously so that all regions of the BM reach their maximum amplitudes at the same moment.

Figure 1(b) shows the waveform of the chirp with compensated delay calculated based on this function. The enhanced delay chirp had the reversed temporal relation of the compensated delay chirp. This enhanced delay chirp was used to investigate whether or not any asymmetry exists due to auditory mechanisms that cancel out the intrinsic cochlear delay between frequency components.

1.2. Auditory Brainstem Response (ABR)

In addition to the psychophysical experiment, we recorded Auditory Brainstem Responses (ABR) from musicians and non-musicians as an electrophysiological response to the asynchronously presented tones. As recent studies demonstrate [14–17], ABR or Frequency Following Response (FFR) provide insight into the underlying mechanisms of auditory function.

A typical ABR consists of five prominent positive peaks that are evoked by a transient sound such as a pulse or a click. These peaks are labelled with Roman numerals I-V. It is assumed that the origin of each peak generally corresponds to the activity generated at each auditory relay nucleus. In this study, the morphological changes of the ABR wave-V, which is generated in the Inferior Colliculus (IC), were investigated.

The ABR wave-V evoked by a stimulus that cancels out cochlear delay is more salient than the ABR wave-V evoked by a stimulus that enhances cochlear delay [12]. It seems that wave-V reflects the synchrony of excitation along the entire BM partition. Don and Eggermont (1978) revealed narrow-band contributions to the ABR from specific portions of the BM [18]. Their results indicated that nearly the entire cochlear partition can contribute to the ABR. Don and Eggermont (1978) also found that the amplitude change of wave-V, as a function of the Central Frequency (CF) of narrow band stimuli, was different from that of waves I and III [18]. For central frequencies above 2 kHz, the wave-V amplitude behaviour is the same as for waves I and III. For CFs below 2 kHz, the wave-V amplitude remained nearly constant throughout the whole CF range, whereas waves I and III showed a rapid drop in amplitude with decreasing CF. In our present experiments, the frequency ranges of stimuli were 0.1 to 10.4 kHz. Thus, investigating the wave-V change appears to be the best measure of the activity of the entire BM. In fact, Dau et al. (2000) reported that clear peaks corresponding to wave I-III could not be observed for stimuli that cancelled out cochlear delay [12].

ABR is an electrophysiological event evoked by the onset of an acoustic stimulus and is generally unaffected by further stimulation. It seems that there is some relationship between the perception of onset asynchrony or multiple onsets, and evoking multiple ABR wave-Vs. Therefore, there is the possibility that the difference in accuracy of onset asynchrony based on musical experience affects the behavior of ABR wave-V. In this study, we observed the relationship between the perception of onset asynchrony and the behavior of ABR wave-Vs.

2. Experiment I

2.1. Thresholds of Synchrony Detection

2.1.1. Apparatus

The experimental sequences were carried out with a PC (dynabook SS RX2, Toshiba). The experimental sequences were carried out with a PC (dynabook SS RX2, Toshiba). The stimuli were generated digitally (48 kHz, 24 bit) and controlled by matrix operation software (Matlab, MathWorks) and diotically presented, i.e., the same sound was presented to both ears of participants through insert earphones (IE8, Sennheiser) driven by an audio interface (UA-101, Roland). The level of each stimulus and the impulse response of the earphones were measured by a sound level meter (type 2250B, Brüel and Kjær) through an ear simulator (type 4157, Brüel and Kjær) with a microphone (type 4192, Brüel and Kjær) and a preamplifier (ZC0032, Brüel and Kjær) (**Fig. 3**). The participants were seated in a soundproof room and responded to the stimuli by pressing a button on a GUI presented on the PC screen.

2.1.2. Stimuli

Three types of sounds were employed as the experimental stimuli: (a) compensated delay chirp, (b) enhanced delay chirp, and (c) pulse. In the compensated delay chirp, the frequency of the sound increased from 0.1 to 10.4 kHz within 10 ms (**Fig. 1**). In the enhanced delay chirp, the frequency of the sound decreased from 10.4 to 0.1 kHz within 10 ms. The chirps were tapered transients at both ends with a raised cosine wave of 0.1 kHz. The pulse was low-pass filtered with a cut-off frequency of 10.4 kHz.

Figure 2 shows frequency and phase responses of the earphone output. As shown in **Fig. 2**, frequency response was flat, below 6 kHz. Phase response was quite linear. It appeared that frequency and phase responses of the earphone output were enough to maintain the waveform of experimental stimuli, because the cochlear delay largely occurs to frequency components below 1000 Hz (**Fig. 1(a)**). The output waveforms of each sound-type were also obtained by convolving the impulse response of the earphone (**Fig. 3**). The waveform and phase of each sound type are closely reproduced.

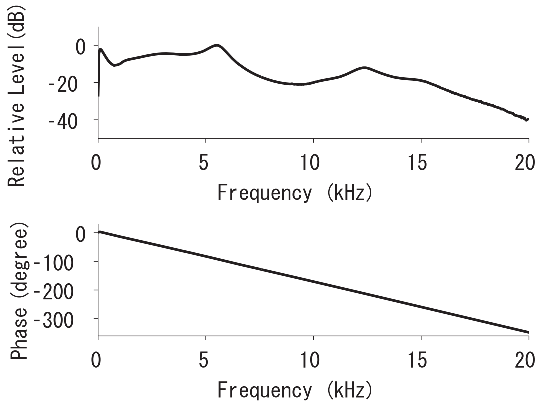


Fig. 2. Frequency response (top) and phase response (bottom) of the earphone output.

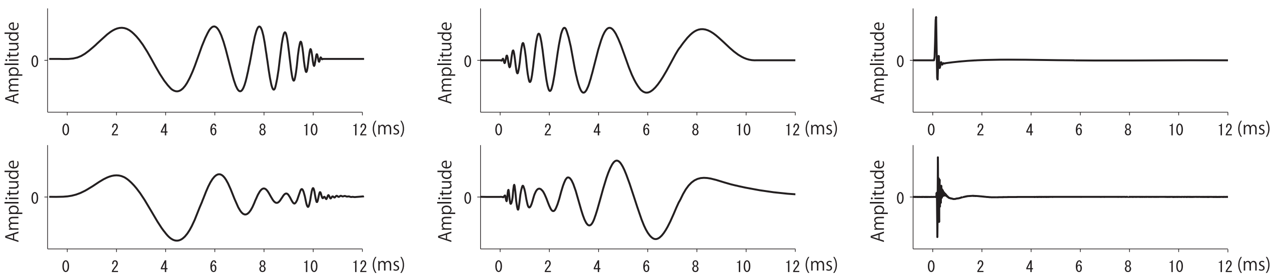
The level of each stimulus presented was approximately 60 dB SPL. They varied from +6 to −6 dB to prevent participants from applying loudness differences as cues for the experimental task. Two sounds of identical sound type were paired to generate synchronous pairs or asynchronous pairs.

2.1.3. Procedure

The experimental task was to detect a synchronous pair in the two-interval, two-alternative forced choice (2I2AFC) procedure. Two pairs of sounds were presented to the participant in one trial (**Fig. 4**). In each trial, one interval contained a synchronous pair and the other interval contained an asynchronous pair. The asynchronous pairs consisted of 12 variations of a temporal gap (0.2, 0.4, 0.7, 1.0, 1.5, 2.3, 2.8, 3.4, 4.1, 5.1, 11.4, or 25.6 ms), spaced roughly logarithmically. The order of the synchronous and asynchronous pairs was randomized across trials to prevent prediction by the participants. The two pairs were separated by a 500-700 ms inter-stimulus interval. The type of the sound was the same within each trial.

Each trial was categorized as one of the following three sound-type conditions: (a) compensated delay, (b) enhanced delay, or (c) pulse. The total number of stimulus-type combinations was 72 (three sound-type conditions, 12 temporal gaps, and two patterns of order of the synchronous pair). The participants repeated each combination 10 times, which brought the total number of trials to 720. All factors (sound type, temporal gap, and presentation order) were randomized and executed as a within-participant factor.

The participants were instructed that both intervals contained two sounds and that the two sounds were synchronous in one interval but asynchronous in the other



(a) compensated (b) enhanced (c) pulse

Fig. 3. The top panels show the input waveforms, and the bottom panels show the output waveforms through the earphones for each sound.

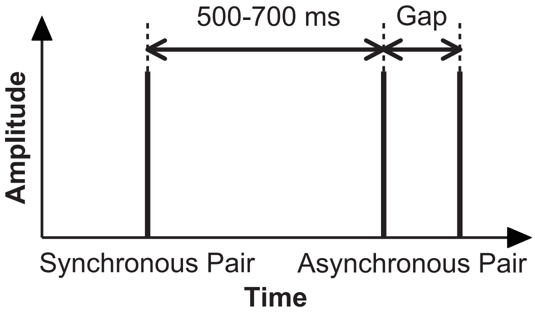


Fig. 4. Example of stimulus presentation by pulse. Here, the first interval contains the "synchronous" pair and should be chosen as "correct." Each vertical line represents one of the three experimental sounds (the compensated delay chirp, the enhanced delay chirp or the pulse).

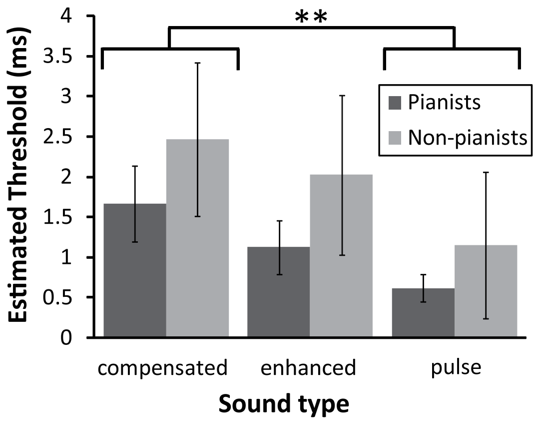


Fig. 5. The average estimated thresholds in ms and the SDs for the conditions of each musical skill and sound type.

interval. They were required to choose the interval containing the synchronous pair. Participants had as many training trials as they wanted and received feedback after each judgment. They were able to take breaks at any time.

Thresholds were estimated from 12 points of psychometric function by fitting a sigmoid function on the results of each participant and computing the temporal gap value corresponding to 75% correct responses.

2.1.4. Participants

Participants were musicians and non-musicians. The musicians were limited to "pianists." Pianists employ performance styles that depend on synchrony, such as chords, arpeggio and playing more than one melody at the same time. Furthermore, they have more opportunities for solo performance compared to players of other instruments, which makes it natural for pianists to control all the tones themselves. Therefore, it is considered that they pay more attention to their performance, particularly to chords. Our previous studies suggest that it is easier to detect onset synchrony for piano tones compared to other instruments, and therefore, it appears that the pianists are making synchrony judgments at a much more precise level [19].

The participants were five pianists (four females and one male) and eight non-pianists (three females and five males) with normal hearing and no history of hearing problems. All pianists had the experience of winning more than one prize in a domestic or foreign competition and have more than 16 years of training (25.4 ± 5.6 years of training). The non-pianists included two people who play the piano or other instruments as amateurs (4.4 ± 6.2 years of training). They were between 20 and 36 years of age and were paid for their participation. This factor was termed "musical skill" and was treated as a between-participant factor.

2.2. Results

Figure 5 shows the average estimated thresholds and the SDs for each musical skill and sound type. A two-way factorial fixed-effect ANOVA was performed where the conditions of musical skill and sound type were treated as the main factors. Musical skill was significant as a main factor ($F(1,37) = 8.61,\ p < .01$). Accuracy of synchrony judgment was significantly higher in pianists than non-pianists. Sound type was also significant as a main factor ($F(2,37) = 7.28,\ p < .01$). There was no interaction between musical skill and sound type.

To investigate detailed differences among the three sound types, we performed an ad-hoc multiple comparison using the Tukey-Kramer HSD test, in which the significant level was adjusted. As a result, only the comparison between the pulse and compensated delay conditions was significant, and the accuracy of the synchrony judgment was significantly lower under the compensated delay condition than under the pulse condition.

2.3. Discussion

Based on the significant main effect of musical skill and the fact that the accuracy of synchrony judgment of pianists was significantly higher than that of non-pianists, we conclude that perceptual accuracy of the pianists increased through the training. Synchrony judgment accuracy under the compensated delay condition was significantly lower than under the pulse condition for both pianists and non-pianists. There was no interaction between musical skill and sound type. There was therefore no significant difference in synchrony judgment accuracy for each sound type that depended on musical skill, which suggests that both pianists and non-pianists have an asymmetric aspect in temporal processing of within the auditory system, and other factors contribute to the more accurate synchrony judgments by pianists.

3. Experiment II

3.1. Measurement of ABR

3.1.1. Apparatus

The experimental sequences were carried out with a PC (dynabook SS RX2, Toshiba). The stimuli were generated digitally (48 kHz, 24 bit) by matrix operation software (Matlab, MathWorks), controlled by audio recording software (SONAR, Cakewalk), and diotically presented to the participants through earphones (IE8, Sennheiser) driven by an audio interface (UA-101, Roland). The level of each stimulus was measured by a sound level meter (type 2250-B, Brüel and Kjær) through an ear simulator (type 4157, Brüel and Kjær) with a microphone (type 4192, Brüel and Kjær) and a preamplifier (ZC0032, Brüel and Kjær).

Silver-silver chloride electrodes were placed at Cz (active) and Fpz (ground) according to the International 10-20 system of electrode placement. The ipsilateral earlobe served as the placement of a reference electrode. The interelectrode impedance was maintained below 5 kΩ. The responses were amplified with a preamplifier (AB601G, Nihon Kohden) using a time constant of 2 seconds and were downloaded to a PC (EPSON, Endeavor NJ5100Pro) at a sampling frequency of 20 kHz through an A/D converter (NI USB-6215, National Instruments).

3.1.2. Stimuli

ABR was measured with a series of chirps with no gap and two chirps with a short gap. The gaps were spaced roughly logarithmically from 0 to 5.1 ms in 8 steps (0, 0.4, 1.0, 2.3, 2.8, 3.4, 4.1, 5.1 ms).

We measured each case using three sound-type conditions (compensated, enhanced and pulse), as in experiment I. The total number of stimulus-type combinations was 24 (three sound-type conditions and 8 temporal gaps).

3.1.3. Procedure

The participants lay on a bed in a soundproof room with electrodes attached to them. They were instructed to keep movement to a minimum and to sleep if possible. The lights were dimmed at the start of measurement.

Stimuli were presented to them for about 4 minutes at a mean repetition rate of about 14 Hz in all 24 stimulus-type conditions. A temporal jitter of ± 4 ms was introduced to minimize the response superimposition from preceding stimuli. As a result, the Inter-Onset Interval (IOI) was equally distributed between 66 and 74 ms. Each trial consisted of 4000 pulses. The total time of measurement was about 2 hours.

The stimulus presentation level was approximately 60 dB SPL.

3.1.4. Participants

The participants were four pianists (three females and one male, 25.3 ± 7.0 years of training) and eight non-pianists (three females and five males, 4.4 ± 6.2 years of training) with normal hearing and no history of hearing problems. All participants also participated in the psychoacoustic experiment.

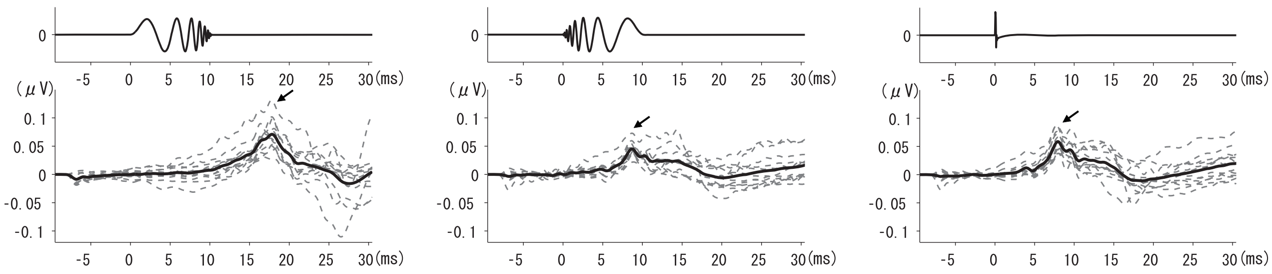
3.2. Results

Figure 6 shows the waveforms of the presented sounds (top) and grand averaged ABR waveforms over all participants (bottom) for each sound-type condition with no gaps. The vertical line indicates wave-V. The baseline was set to be zero on average before stimulus onset.

Saliency and latency of wave-Vs for each sound type were clearly different (**Fig. 6**). The waveform of wave-V for the compensated delay condition appeared to be the most salient, and the enhanced delay condition had the gentlest peak. These results correspond to those of Dau et al. (2000) [12].

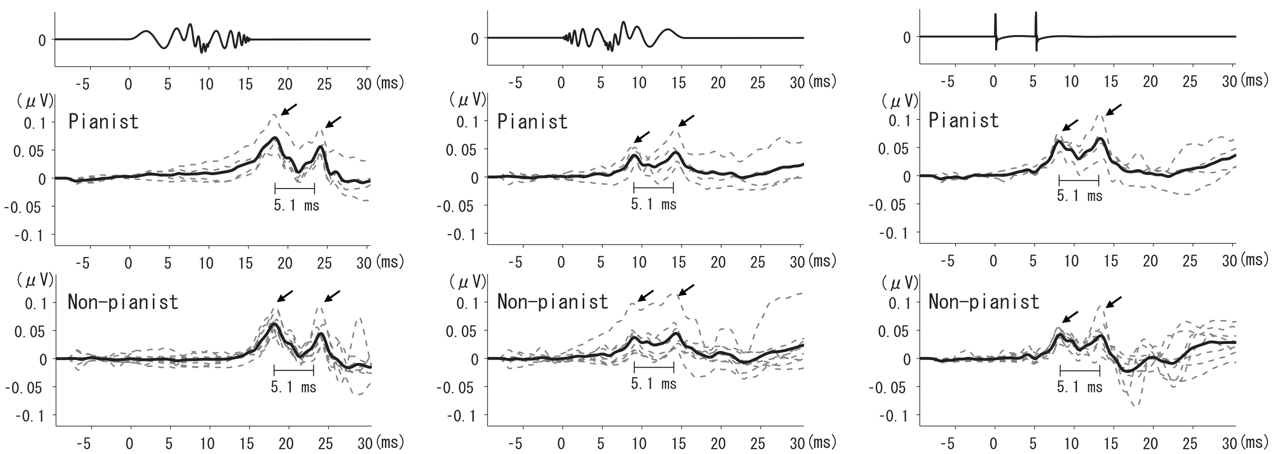
For the latency of the wave-V, a two-way factorial fixed-effect ANOVA was performed, in which musical skill and sound type were treated as the main factors. The sound type was significant as a main factor ($F(2,35) = 3956.91$, $p < .01$), and the Tukey-Kramer HSD test showed that there were significant differences among all sound-type conditions. We found no significant difference between the two musical skill levels and no significant interaction between musical skill and sound type. The latency of wave-V was longest in the compensated delay condition, followed by the enhanced delay condition and the pulse condition.

For stimuli with 3.4, 4.1 and 5.1 ms gaps, the intervals between the two wave-Vs (one is considered to be evoked by the first chirp and another by the second chirp) were measured. However, for stimuli with 0.4 to 2.8 ms gaps, the second wave-Vs did not show clear peaks and were not clearly distinguishable from the first wave-V. **Fig. 7** shows the waveforms of the presented sounds (top) and the grand averaged ABR waveforms over four pianists (middle) and eight non-pianists (bottom) for each sound-type condition with 5.1 ms gaps. Under these conditions, a time difference between the gap of physical stimuli of 5.1 ms and the interval between two wave-Vs was discernible.



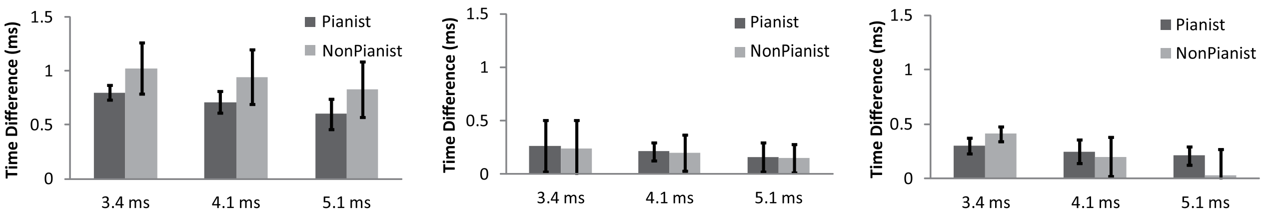
(a) compensated (b) enhanced (c) pulse

Fig. 6. Black lines indicate waveforms of presented sounds (top) and grand averaged ABR waveforms over all participants (bottom) for each sound-type condition with no gap. The grey broken lines indicate waveforms of each participant (bottom). The arrow indicates wave-V.



(a) compensated (b) enhanced (c) pulse

Fig. 7. Black lines indicate waveforms of the presented sounds (top) and grand averaged ABR waveforms over four pianists (middle) and eight non-pianists (bottom), for each sound-type condition with 5.1 ms gap. The grey broken lines indicate waveforms for each participant (middle, bottom). The arrow indicates wave-V. The short horizontal line indicates time interval of 5.1 ms.



(a) compensated (b) enhanced (c) pulse

Fig. 8. The average time differences between the physical stimuli gap and the interval between two wave-Vs in ms and the SDs for each musical skill, sound type and gap.

We therefore used the time difference between the physical stimuli gap and the interval between two wave-Vs as an index to observe ABR behaviour. This is considered to be the index that should not be affected by the individual difference of latency.

We performed a three-way factorial fixed-effect ANOVA for the time difference between the physical stimuli gap and the interval between the two wave-Vs where musical skill, sound type and gap were treated as the main factors. **Fig. 8** shows the time difference between the physical stimuli gap and the interval between the two wave-Vs. The gap conditions were limited to 3.4, 4.1 and 5.1 ms, the conditions under which the first and second wave-Vs could be clearly distinguished.

Sound type ($F(2,107) = 106.90,\ p < .01)$) and gap ($F(2,107) = 524.44,\ p < .01$) were significant as main factors. There was an interaction between musical skill and sound type ($F(2,107) = 4.94,\ p < .01$). A Tukey-Kramer HSD test on these results showed that for sound type, the time difference between the physical stimuli gap and the interval between the two wave-Vs under the compensated delay condition was significantly longer than under the other two conditions. For the gap, the comparison between the 5.1 ms gap and 3.4 ms was significant, and the time difference of the 5.1 ms gap was significantly shorter than the 3.4 ms gap. For the interaction between musical skill and sound type under the compensated delay condition, the time difference between the physical stimuli gap and the interval between the two wave-Vs of non-pianists was significantly longer than that of pianists.

Though these results were obtained by the limited number of pianists, it is also a fact that there were statistically significant differences.

3.3. Discussion

For the saliency and the latency of wave-V, there were no differences between pianists and non-pianists, suggesting that training does not affect characteristics of cochlear delay.

Because the waveform of the wave-V evoked by the compensated delay chirp was more salient than the pulse and the enhanced delay chirps, it seems that the wave-V reflects more synchronous excitation of the entire BM. Auditory nerve fibres tend to fire at a particular phase during each cycle of vibration of the BM [20], which is called "phase-locking." The firing of auditory nerve fibres evoked by the compensated delay chirp is considered to start more simultaneously than that of the pulse and that of the enhanced delay chirps. Don and Eggermont (1978) revealed that nearly the entire BM partition can contribute to the ABR [18]. These results suggest that a more salient wave-V reflects more synchronous excitation of the whole BM.

The latency of the compensated delay chirp is significantly longer than the other stimuli, followed by the enhanced delay chirp and the pulse. As shown in **Fig. 6**, the latency of the compensated delay chirp is notably longer than that of the other two sound types; however, the latency of the enhanced delay chirp and the pulse are comparable. In fact, the average latency of the compensated delay chirp, the enhanced delay chirp and the pulse over all participants was approximately 17.3, 8.3 and 7.5 ms, respectively. The duration of the compensated delay chirp was approximately 10.3 ms, suggesting that the interval between the offset of the compensated delay chirps and the evoked wave-V was approximately 7.0 ms. The compensated delay chirp was designed to induce simultaneous excitation along the entire BM. Considering the fact that the latency of wave-V was about 8 ms from the point of sound input, it is generally assumed that the wave-V was evoked in approximately 7 ms after excitation of the BM [21]. The latency of the enhanced delay chirp was significantly longer than the pulse, suggesting a relationship between the spread of the stimulus and enhancement of excitation at BM.

For latency, there was also no difference between pianists and non-pianists. This result suggests cochlear delay is due to mechanical characteristics of the BM and the timing of the BM excitation does not vary with training or influence synchrony judgment accuracy.

For the 3.4, 4.1 and 5.1 ms gap conditions, the time difference between the physical stimuli gap and the interval between the two evoked wave-Vs for the compensated delay condition was significantly longer than for the enhanced delay and pulse conditions. Additionally, a comparison between the 5.1 ms gap and 3.4 ms showed significant differences. It appeared that longer gaps between physical stimuli resulted in shorter time differences between physical stimuli gaps and intervals between the two wave-Vs. This result suggests that the shorter time difference between the physical stimuli gap and the interval between two wave-Vs results in more accurate synchrony judgment as measured by the psychophysical experiment.

This agrees with the observation that, under the compensated delay condition, the time difference between the physical stimuli gap and the interval between the two wave-Vs of non-pianists was significantly longer than that of pianists. In the wave-V of pianists, there may be less interference from the wave-V of secondary sounds evoked after a subtle gap, compared to that of non-pianists. Although still at the hypothesis stage, there may be increased firing synchrony in the auditory nerve. In other words, the non-pianists are more easily affected by secondary sounds than the pianists because firing in the auditory nerve of non-pianists exhibits a wider temporal spread. The attentive listening required for musical training may affect the neural encoding of sound stimuli. This result supports the conclusion of Bidelman [14] that long-term experience with auditory signals induces experience-dependent neural plasticity, the ability for the nerve cells to change or modulate their function [21]. The pianists who participated in this experiment started their piano training ages between 3 and 5 years old, and they have more than 16 years of continuous training. It is therefore possible that neural plasticity has been induced in these subjects by continuous training from a very young age, when their brains were still in the developmental stages. According to previous studies [16, 17], both a long period of training and training during development stages are important.

We observed no such differences in the enhanced delay or pulse condition. For the enhanced delay and pulse condition, this may suggest that there was already a time difference between the pianists and non-pianists, which occurred during shorter gaps than the compensated delay condition. There is, therefore, the possibility that because these phenomena were evoked during a very short gap, they are difficult to distinguish from the second wave-V.

In the enhanced delay and pulse conditions, the physical stimulus gap was almost equal to the interval between the two wave-Vs. This result suggests that under the enhanced delay and pulse conditions, the interference of the first wave-V with the second wave-V was less than that observed under the compensated delay condition. Because higher frequencies vibrate the basal side of the BM, in the pulse and enhanced delay conditions, these vibrations by the higher frequencies occur without coincident vibration at lower frequencies. In the compensated delay condition, however, the lower frequencies pass through the basal side first in order to travel to the corresponding apical side. This vibration by the lower frequencies might interfere with vibrations on the basal side and partially mask the cue given by the higher frequencies. This interference may influence the time difference between the physical stimuli gap and the interval between two wave-Vs evoked by the compensated delay chirps than that of other two sound type conditions. Thus, the vibration of the basal side of the BM evoked by higher frequencies is an effective cue for synchrony judgments.

In our previous study [11], BM motion was simulated by the auditory peripheral model [22] using the same three types of sound that were used in this experiment. In this auditory peripheral model, we simulated the vibration of each channel after dividing the basilar membrane of cochlea into 100 channels ranging from the apical side to the basal side. The results showed that asynchronous stimuli tended to produce two clear peaks in the high-frequency channels but not in low-frequency channels. Therefore, the information provided by the high-frequency channels might be more effective as a cue to make synchrony judgments than those provided by the lower-frequency channels.

4. Conclusions

Synchrony judgment is an important ability when either playing or listening to music. We have performed a series of phychoacoustic and physiological experiments to investigate if long-term musical training would improve accuracy of synchrony judgments, and also if ABR would change accordingly. In addition, we investigated whether there are differences in the asymmetric aspects of temporal processing in the human auditory system between pianists and non-pianists.

The results of the psychoacoustic experiments showed that the accuracy of synchrony judgments of pianists was higher than that of non-pianists on all types of sounds. We therefore concluded that the ability related to the temporal tone perception was improved through piano training.

The time difference between the physical stimuli gap and the interval between the two wave-Vs of non-pianists was significantly longer than that of pianists. In the case of the wave-V of the pianists, we propose that there is less interference in the wave-V evoked by secondary sounds after a subtle gap, compared to that of non-pianists. Therefore, the synchrony of firing of the auditory nerve may be increased by training. Our results support the idea that long-term experience with auditory signals induces experience-dependent plasticity of the brainstem [23].

However, for the saliency and the latency of wave-V, there were no differences between pianists and non-pianists. This suggests that asymmetric aspects of temporal processing in the human auditory system do not change along with such experience-dependent plasticity.

Acknowledgements

This work was supported by a Grant-in-Aid for Young Scientists (B), No.22730592 and No.23730715 and Grant-in-Aid for Scientific Research (B) No.21330170, Japan Society for the Promotion of Science.

References:

[1] A. S. Bregman, "Auditory scene analysis," MIT Press, 1990.
[2] C. J. Darwin, "Perceiving vowels in the presence of another sound: constraints on formant perception," J. Acoust Soc Am, Vol.76, No.6, pp. 1636-1647, 1984.
[3] C. J. Darwin and V. Ciocca, "Grouping in pitch perception: effects of onset asynchrony and ear of presentation of a mistuned component," J. Acoust Soc Am, Vol.91, Vol.6, pp. 3381-3390, 1992.
[4] R. W. Hukin and C. J. Darwin, "Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification," Percept Psychophys, Vol.57, No.2, pp. 191-196, 1995.
[5] J. A. Sloboda, "Generative processes in music: The psychology of performance, improvization, and composition," Oxford University Press, Oxford, pp. 72-90, 1988.
[6] R. A. Rasch, "The perception of simultaneous notes such as in polyphonic music," Acustica, Vol.40, pp. 22-33, 1978.
[7] S. Sadie and J. Tyrrell, "The New Grove Dictionary of Music and Musicians," Oxford University Press, USA, 2004.
[8] G. von Békésy, "Experiments in hearing," McGraw-Hill, New York, 1960.
[9] S. Uppenkamp, S. Fobel, and R. D. Patterson, "The effects of temporal asymmetry on the detection and perception of short chirps," Hearing Research, Vol.158, No.1-2, pp. 71-83, 2001.
[10] E. Aiba and M. Tsuzaki, "Perceptual judgment in synchronization of two complex tones : Relation to the cochlear delays," Acoustical science and technology, Vol.28, No.5, pp. 357-359, 2007.
[11] E. Aiba, M. Tsuzaki, S. Tanaka, and M. Unoki, "Judgment of perceptual synchrony between two pulses and verification of its relation to cochlear delay by an auditory model," Japanese Psychological Research, Vol.50, No.4, pp. 204-213, 2008.
[12] T. Dau, O. Wegner, V. Mellert, and B. Kollmeier, "Auditory brainstem responses (ABR) with optimized chirp signals compensating basilar membrane dispersion," J. Acoust Soc Am, Vol.107, No.3, pp. 1530-1540, 2000.
[13] E. de Boer, "Auditory physics: Physical principles in hearing theory I," Physics Report, Vol.62, pp. 87-174, 1980.
[14] G. M. Bidelman and A. Krishnan, "Effects of reverberation on brainstem representation of speech in musicians and non-musicians," Brain Research, Vol.1355, pp. 112-125, 2010.
[15] G. M. Bidelman, A. Krishnan, and J. T. Gandour, "Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch," Eur J. Neurosci, Vol.33, No.3, pp. 530-538, 2011.
[16] G. Musacchia, D. Strait, and N. Kraus, "Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians," Hearing Research, Vol.241, No.1-2, pp. 34-42, 2008.
[17] D. L. Strait, N. Kraus, E. Skoe, and R. Ashley, "Musical experience and neural efficiency? effects of training on subcortical processing of vocal expressions of emotion," European J. of Neuroscience, Vol.29, No.3, pp. 661-668, 2009.
[18] M. Don and J. J. Eggermont, "Analysis of the click-evoked brainstem potentials in man using high-pass noise masking," J. Acoust Soc Am, Vol.63, No.4, pp. 1084-1092, 1978.
[19] E. Aiba, ""Unification" and "Separation" of Overlapped Sounds," Doctral Thesis, Graduate School of Kyoto City University of Arts, 2009.
[20] J. O. Pickles, "An introduction to the physiology of hearing," Emerald (3rd), United Kingdom, 2008.
[21] A. R. Møller, "Hearing : anatomy, physiology, and disorders of the auditory system," Elsevier Academic, Amsterdam ; Oxford (2nd), 2006.
[22] Y. Murakami and M. Unoki, "A study on the input-output function of a nonlinear cochlear transmission-line model with the function of an outer hair cells model," J. Acoust Soc Am (151st Meeting of the Acoustical Society of America, New Orleans, Louisiana, USA), Vol.122, p. 3061, 2007.
[23] T. Tzounopoulos and N. Kraus, "Learning to encode timing: mechanisms of plasticity in the auditory brainstem," Neuron, Vol.62, No.4, pp. 463-469, 2009.

Name:
Eriko Aiba

Affiliation:
School of Science and Technology, Kwansei Gakuin University

Address:
2-1 Gakuen, Sanda, Hyogo 669-1337, Japan

Brief Biographical History:

2004 B.A. in Piano Performance, Kyoto City University of Arts
2006 M.A. in Musicology, Graduate School of Music, Kyoto City University of Arts
2009 Ph.D. in Musicology, Graduate School of Music, Kyoto City University of Arts
2009- Post-Doctoral Fellow at Kwansei Gakuin University

Main Works:

• "Judgment of perceptual synchrony between two pulses and verification of its relation to cochlear delay by an auditory model," Japanese Psychological Research, Vol.50, pp. 204-213, Nov. 2008.

Membership in Academic Societies:

• The Acoustical Society of Japan (ASJ)
• The Japanese Society for Music Perception and Cognition (JSMPC)
• The Society for Neuroscience

Name:
Koji Kazai

Affiliation:
Graduate School of Science and Technology, Kwansei Gakuin University

Address:
2-1 Gakuen, Sanda, Hyogo 669-1337, Japan

Brief Biographical History:

1993 B.A. in Psychology, Kwansei Gakuin University
1995 M.A. in Psychology, Kwansei Gakuin University
2003 Ph.D. in Psychology, Kwansei Gakuin University
2003- Research Fellow at Kwansei Gakuin University

Main Works:

• "Haptic vibration improves the perceived quality of drum sounds," Trans. of Japan Society of Kansei Engineering, Vol.9, pp. 591-600, Sep. 2010.

Membership in Academic Societies:

• Japanese Psychological Association (JPA)
• Japanese Society for Physiological Psychology and Psychophysiology (JSPP)
• Japan Society of Kansei Engineering (JSKE)

Name:
Takayuki Shimotomai

Affiliation:
Brain Science Institute, Tamagawa University

Address:
6-1-1 Tamagawa-gakuen, Machida, Tokyo 194-8610, Japan

Brief Biographical History:

2004- Research Institute, National Rehabilitation Center
2007- The Department of Informatics, Kwansei Gakuin University
2010- Brain Science Institute, Tamagawa University

Main Works:

• "Model-Based Analysis of Verb Meaning Inference in Infant, Cognitive Studies," Vol.11, No.4, pp. 319-332, 2004.

Membership in Academic Societies:

• Japanese Neural Network Society (JNNS)
• The Society for Neuroscience
• The Robotics Society of Japan (RSJ)

Name:
Toshie Matsui

Affiliation:
Head and Neck Surgery, Department of Otorhinolaryngology, Nara Medical University

Address:
840 Shijo-cho, Kashihara, Nara 634-8522, Japan

Brief Biographical History:

2003 M.A. in Piano Performance, Graduate School of Music, Kyoto City University of Arts
2010 Ph.D. in Musicology, Graduate School of Music, Kyoto City University of Arts
2010- Post-Doctoral Fellow at Kwansei Gakuin University
2011- Post-Doctoral Fellow at Nara Medical University

Main Works:

• "Functional difference between the tonotopic and periodic information in recognition of transposed melodies: How do local cues affect global features?" Acoustical Science and Technology, Vol.29, No.5, pp. 309-319, 2008.

Membership in Academic Societies:

• The Acoustical Society of Japan (ASJ)
• The Japanses Psychonomic Society
• The Oto-Rhino-Laryngological Society of Japan

Name:
Minoru Tsuzaki

Affiliation:
Professor, Faculty of Music, Kyoto City University of Arts

Address:
13-6 Kutsukake-cho, Oe, Nishikyo-ku, Kyoto 610-1197, Japan

Brief Biographical History:
1982- Started his academic career as an Assistant Professor in psychology at Niigata University
1988- Joined Advanced Telecommunication Research Laboratories
2004- Kyoto City University of Arts

Main Works:
• "Effects of the preceding scale on melodic interval judgment in terms of equality and size," Music Perception, Vol.9, pp. 47-70, 1991.
• "Shrinkage of perceived tonal duration produced by extra sounds: effects of spectral density, temporal position, and transition direction," Perception, Vol.29, pp. 989-1004, 2000.
• "Perception of size modulated vowel sequence: Can we normalize the size of continuously changing vocal tract?" Acoustical Science and Technology, Vol.30, pp. 83-88, 2009.

Membership in Academic Societies:
• Acoustical Society of Japan (ASJ)
• Acoustical Society of America (ASA)
• Japanese Psychological Association (JPA)
• Japanese Psychonomic Society (JPS)
• Japanese Society of Music Perception and Cognition (JSMPC)

Name:
Noriko Nagata

Affiliation:
Professor, School of Science and Technology, Kwansei Gakuin University

Address:
2-1 Gakuen, Sanda, Hyogo 669-1337, Japan

Brief Biographical History:
1983- Joined Mitsubishi Electric Corp.
2003- Associate Professor, Kwansei Gakuin University
2007- Professor, Kwansei Gakuin University

Main Works:
• "Modeling and Visualization for a Pearl-Quality Evaluation Simulator," IEEE Trans. Visualization and Computer Graphics, Vol.3, No.4, pp. 307-315, Oct. 1997.

Membership in Academic Societies:
• The Institute of Electrical and Electronics Engineers (IEEE)
• Association for Computing Machinery (ACM)
• Information Processing Society of Japan (IPSJ)