**ABSTRACT**

The model for pitch assumes that pitch is based on the periodicity in the neural activities after the cochlear filtering. One could argue that the auditory system "uses" the pitch as cue for stream segregation. A question, however, would whether pitch is a cause or an end of such grouping. We investigated the case where two pulse trains with an identical periodicity are added with variable temporal disparities. The second pulse train with the identical IPI was added with various phase delays. When the phase delay was 50%, the pitch raised by an octave. This impression of the octave shift appeared to be continuous as a function of the degree of the phase delay except for a hump was observed at 25% point. The auditory model could not provide any corresponding peak in the time interval histogram of the neural activities. Another series of experiments by the authors indicated that aged absolute pitch possessors tended to perceive pitches higher than young AP possessors. An additional experiment using experimental sounds indicated that similar results could be obtained only for sounds having temporal information in the lower order region.

**Keywords:** Pitch, Temporal Information, Aging  
**I-INCE Classification of Subjects Number(s):** 76.9

1. **INTRODUCTION**

When we consider auditory noise, it is assumed that there might always be a signal as its conceptual counterpart. Noise is assumed to be harmful to detect the existence of any signal. Signals are, in its broad sense, auditory events of some importance, and they are usually assumed to have a certain regularity, i.e., periodicity. Pitch percepts have been considered to be mental correlates to the periodic sounds.

The core of this conference is the issue on noise. The concept of noises, however, should always be considered in reference to the concept of signals, and it seems to be implied that signals possess periodicity, or quasi-periodicity, in most of cases. This implication (or tacit assumption) has a physical rationale. For example, the vowel categories are determined by the shape of the vocal tract in human vocal communication, and theoretically an impulse response evoked by a single pulse can supply sufficient informations. This single impulse response, however, can critically suffer from an accidental co-occurrence of another impulsive sound. Therefore, repeating impulse responses periodically can overcome this accidental mixture as well as can make the vowels to pop out against random background noises. In this sense, the periodicity works as a channel identifier in an auditory communication.

Needless to mention, the auditory theories have been developed to explain how our auditory system can extract the periodicity of auditory signals with such high precision (1, 2). The periodicity is temporal and can be described in terms of time, and the pitch phenomena are occurring in the order of millisecond, which could intuitively appear too fast to follow by our daily time perception. Therefore, some pioneers in the auditory science, such as von Helmholtz, tended to seek for a
mechanism which can represent the periodicity of in terms of the resonance characteristics of the vibrating bodies (3). While this approach have lead to successful research ends in the mechanical properties of the cochlea, there exist a couple of pitch phenomena which still require temporal information, such as the pitch of missing fundamentals (4). The finding of the phase locking in the auditory nerve activities has supplied the physiological reality of the temporal coding of the physical periodicity (5).

Those studies are summarized and have provided a conceptual frame work for the main stream of the current auditory models where a certain mechanism is assumed to work to make time interval histograms in the phase locked neural activities obtained for the output the frequency analysis by the cochlear filter bank (6, 7). Thus, pitch has been one of the important target for the auditory theory to explain.

Most of people would agree that the pitch reflect the periodicity of the sounds. The meaning of the word, “sounds”, has two aspects; (a) a sound source in the physical domain, and (b) a subjective impression in the mental domain. This leaves a room for the circular reasoning. If two sound sources emit pulse trains with different periodicity, it is rather difficult to find any fixed periodicity between each consecutive pulses in the sound mixture obtained by a simple addition of two wave forms. To find each periodicity, one must segregate the sound mixture appropriately. It is, however, the differences in pitch in this situation that signify the sound source (8). It seems rather a tricky question whether pitch is the end of perceptual segregation, or the cause of the perceptual segregation.

In the first part of the current paper, the pitch perception under the condition of adding two identical pulse trains are discussed with some experimental findings.

2. EXPERIMENT 1: SPAIPT – Addition of Two Identical Pulse Trains with Phase Differences – (Impression of Octave Shift)

2.1. Purpose

In Experiment 1, pitch perception for several variations of adding two pulse trains with an identical periodicity was investigated. The variation was achieved by changing the phase difference between the two pulse trains relative to the period of the pulse train. One may naively suppose that we would still hear the identical pitch even if two pulse trains simultaneously exist as far as both have the identical periodicity. With another step of consideration, however, one could also notice an existence of a special case where a second pulse train is added on the first one with a phase delay which is exactly the half of the original periodicity. The resultant wave form becomes a simple pulse train whose period is just the half of the original period, whose pitch should be higher than that of the original by an octave.

A question for the Experiment 1 was how pitch perception changed for the stimuli between the two cases, or where the pitch raised by an octave relative to the original one. To address this question, a series of stimuli, called SPAIPTs, “Split Phase Addition of Identical Pulse Trains”, were prepared. In SPAIPT, two pulse trains with the identical period were added with a phase delay defined in terms of the cycle of the original period.

2.2. Method

2.2.1. Stimulus

PSAIPTs are a series of stimuli which are obtained by adding two pulse trains with an identical periodicity. 

![Figure 1 – Relations between the addition of two pulse trains with a relative phase delay: (a) 0%; (b) 25%; (c) 50%.

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>25%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
periodicity with a various temporal disparities. If one defines the temporal disparities in terms of the percentage of the “original” period, a PSAIPT with 0 % disparity ($\Delta d$) is a pulse train whose periodicity is equal to the original one and whose amplitude is doubled as depicted in Fig. 1(a). At the other end, a PSAIPT with 50 % disparity is a pulse train whose periodicity is a half of the original one as depicted in Fig. 1(c). A stimulus continuum of PSAIPTs ranging from 0 to 50 % disparity could produce a perceptual continuum in which pitch would raise by an octave at a certain point.

SPAIPTs were generated using a pulse train whose period was 5 ms. Another pulse train with the identical amplitude and period was added onto the first pulse train with a delay, ($\Delta d + 380$) ms. The ratio of $\Delta d$ against the fundamental period was ranged from 0% to 50% by 2.5% step. To help listeners to have the unambiguous image of the original pitch, a pure tone whose periodicity was also 5 ms was presented as a reference stimulus. The second section of the combined pulse train lasted for 400 ms.

2.2.2. Procedure

Each listener was required to evaluate the prominence of upward pitch shift by an octave for the second combined pulse train using a five point scale: (1) no octave shift; (5) obvious octave shift. Each experimental session had 10 blocks of 21 trials. In each block, 21 variation of SPAIPTs with different $\Delta d$ s were presented once respectively in a random order. Thus, each participant provided 210 responses in total.

2.2.3. Participants

Twenty undergraduates of Kyoto City University of Arts participated in the experiment. None of them had any serious hearing deficit. All of them were members of Faculty of Music, and had a clear image of the octave difference in pitch.

2.3. Results

Raw evaluation scores of the octave shift were averaged over the repetitions as well as the participants, and were plotted as a function of $\Delta d$ as shown in Fig. 2. The prominence of the octave shift generally tended to increase as the $\Delta d$ increased. In addition, a peculiar point appeared at 25% of the relative delay point. At this point, the averaged score exceptionally higher than what would be predicted from the general trend.

2.4. Discussion

The results suggest two aspects of the phenomenon. The first is depicted by the general increasing trend by ignoring the 25% point observed in Fig. 2, indicating that the impression of the octave rise of pitch was enhanced as a function of the $\Delta d$ s in a continuous manner. This suggested that the information conveyed by the phase difference, i.e., $\Delta d$ s, could be reasonably transmitted to

![Figure 2 – Perceptual strength of the upward octave pitch shift as a function of the relative phase delay between the two pulse trains.](image-url)
the auditory system. It should, however, be noted that this continuous shift of the pitch did not correspond to the normal pitch change associated with the change in the fundamental frequency. The introspections by the participants suggested that the pitch chroma did not change. Some previous studies reported a similar gradual pitch change with no chroma change.

The second aspect is depicted by the peculiar deviation from the general trend at 25% point. The impression of the octave rise at 25% point appeared to be higher than its neighbors. One might argue that this would not be so anomalous because it is exactly a quarter of the original period. Although it is not an octave, it is two octaves. Therefore, it would be likely for the participant to hear any impression of tones higher than the original pitch.

Then, a question would be whether this raised impression at 25% point really reflected the pitch impression higher by an octave, or it simply reflected the pitch impression higher by two octave. Although the participant were instructed to evaluate the strength of the pitch shift by an octave, they had no way to express another impression even if they heard a pitch higher by two octave. Therefore, Experiment 2 was designed.

3. EXPERIMENT 2: SPAIPT (Pitch Matching)

3.1. Purpose

The first purpose of Experiment 1 was to check which impression of the pitch shift was dominant for the 25% point, an octave, or two octaves. To address this, a pitch matching paradigm was used. Participants were allowed to hear several comparison candidates including a tone higher by two octaves as well as one higher by an octave. If the comparison candidate higher by two octaves were chosen more than the one higher by an octave, the augmented impression at 25% point in Experiment 1 simply reflect two octave shift corresponding to a quarter of the original period.

The second purpose of Experiment 2 was to investigate the validity of the “old-plus-new” heuristics, which is one of the heuristics of the auditory scene analysis proposed by Bregman (9). It reflects our perceptual trait that an abrupt addition of sound energy is likely to be perceived as a new different event even when the added “new” sound is a exact copy of the preceding “old” sound. To address this issue, two version of SPAIPTs were prepared, i.e., SPAIPTs with/without a leading segment.

3.2. Method

3.2.1. Stimulus

As to the stimuli with the leading segment, the same SPAIPTs were generated in the same manner as in Experiment 1. The second pulse train was added with 340 ms delay of the first, original pulse train. Thus, the leading segment lasted 340 ms, and this segment was assumed to form an “old” auditory stream in the framework of the “old-plus-new” heuristics. The duration of the second segment was 400 ms where the two pulse trains were added. The $\Delta d$ ratios against the base period

<table>
<thead>
<tr>
<th>Relative Phase Difference</th>
<th>Interval $\Delta d$ (Frequency)</th>
<th>Complement of $\Delta d$ (Frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>$\infty$ (200 Hz)</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>0.25 ms (4000 Hz)</td>
<td>4.75 ms (211 Hz)</td>
</tr>
<tr>
<td>10%</td>
<td>0.5 ms (2000 Hz)</td>
<td>4.5 ms (222 Hz)</td>
</tr>
<tr>
<td>12.5%</td>
<td>0.625 ms (1600 Hz)</td>
<td>4.375 ms (229 Hz)</td>
</tr>
<tr>
<td>15%</td>
<td>0.75 ms (1333 Hz)</td>
<td>4.25 ms (235 Hz)</td>
</tr>
<tr>
<td>20%</td>
<td>1 ms (1000 Hz)</td>
<td>4 ms (250 Hz)</td>
</tr>
<tr>
<td>25%</td>
<td>1.25 ms (800 Hz)</td>
<td>3.75 ms (267 Hz)</td>
</tr>
<tr>
<td>30%</td>
<td>1.5 ms (667 Hz)</td>
<td>3.5 ms (286 Hz)</td>
</tr>
<tr>
<td>45%</td>
<td>2.25 ms (444 Hz)</td>
<td>2.75 ms (364 Hz)</td>
</tr>
<tr>
<td>50%</td>
<td>2.5 ms (400 Hz)</td>
<td>- - -</td>
</tr>
</tbody>
</table>
was 5, 10, 12.5, 15, 20, 25, 30, 45, and 50 %. The reference stimulus was changed to a pulse train from a pure tone with the identical period of the original pulse train, i.e., 5 ms.

The stimuli without the leading segment consisted of only the second 400 ms of the stimuli with the leading segment.

Nineteen types of pulse trains were prepared as comparison candidates for pitch matching. Table 1 shows their periodicities (as well as frequencies). They were simple periodic pulse trains. As the SPAIPTs contained the interval corresponding to $\Delta d$ as well as its complement, a series of pulse trains with those periodicity were prepared, and the candidate which could be higher than the original by two octave was covered in this series. The series also reasonably covered a pitch continuum with different chroma ranging from the original one to the one higher by an octave.

3.2.2. Participant

Twenty undergraduates of Kyoto City University of Arts participated in the experiment. Eighteen of them had participated in Experiment 1. None of them had any serious hearing deficit. All of them were members of Faculty of Music, and had a clear image of the octave difference in pitch.

3.2.3. Procedure

Each participant was required to choose a candidate whose pitch matches most to the test SPAIPT. He/She was allowed to listen to the SPAIPT and matching candidates as many time as he/she wanted to until he/she reached the final decision.

In one experimental session, the test stimuli were fixed to either of the SPAIPTs with/without the leading segment. Each participant was tested both conditions once, and the order of the tests was counter-balanced between the participants. Each session consisted of 10 blocks, in which nine pitch matchings were required to each of nine $\Delta d$ s. Thus, each participant made 180 pitch matchings in total.

3.3. Results

Proportion of the choice by 20 listeners is depicted as the size of discs in Fig. 3 and 4, for the with-preceding condition, and without-preceding condition, respectively. The abscissa represent the ratio of $\Delta d$ against the base period, and the ordinate represent the matched fundamental frequency, i.e., the inverse of the periodicity of the comparison candidates.

In both conditions, the upward match by an octave naturally dominated for the 50% phase delays, and so did it for the 45% cases. The viewgraph also indicates that the primary choice was the original pitch for the 25% case, but that there are some cases of the match to the pitch an octave higher than the original. It should be noted that the matching to the pitch two octave high did occur just once for the with-preceding condition, and did not occur at all for the without-preceding condition. The tendency of one octave pitch shift was less prominent for the 22.5% and the 27.5% cases than the 25% case.

Fig. 3 and 4 also showed that the pitch match to the comparison other than the original and its octave family did seldom occur.

Pitch matching to the comparison candidates other than the original pitch was more likely to occur in the with-preceding condition than in the without-preceding condition.

3.4. Discussion

The probability of the pitch matching to the comparison higher by an octave increased at the 25% point compared to its neighbors, although the majority of the pitch matching was settled to the original pitch. The matching to the comparison higher by two octave was quite exceptional. Actually, only one participant out of 20 made this choice.

Therefore, being combined with the results of Experiment 1, the 25% phase delay was likely to bias the pitch judgement to one octave higher, and it is unlikely for the listeners to perceive a pitch two octaves higher than the original and to misunderstand that it was the pitch one octave higher than the original.

Another important aspect of the results was the response choices were limited to the comparison which had the original pitch and its octave families. In Experiment 2, the comparison candidate having pitches between the original and its octave were provided. Nevertheless, they were not
chosen. Although it may superficially appear inconsistent with the continuous manner observed in Fig. 2, these results characteristically depict the pitch continuum with a fixed chroma.

4. MODEL SIMULATION

It is a well-known conceptual model that the perceptual space for pitch has a multidimensional structure (10). While the apparent pitch rises continuously as the fundamental frequency of a tone increases, it “returns” to the start point when the fundamental frequency becomes exactly doubled, i.e., an octave relationship sharing the same pitch chroma like C, D, E, and so on, in musical term.

Such relations can be represented in a three dimensional space with a helical structure as depicted in Fig. 5. The usual pitch shift achieved by the change in the fundamental frequency tracks the helical path. When the fundamental frequency is doubled, it shares the same position on the base plane. If the pitch change tracks the path indicated by dash lines with arrows in Fig. 5, a pitch change with a fixed chroma can be achieved.

Some previous studies reported that this sort of pitch shifts was achievable either by attenuating the odd harmonics or by applying resonant scaling to every other impulse responses of harmonic complexes like vowels (11, 12). Modern auditory models calculating time interval histograms of the neural activities after a cochlear bank process could reasonably capture such pitch transitions.

Stabilized Auditory Images (SAIs) (13) for some of the SPAIPTs are depicted in Fig. 6. The panel (a) depict the simulation results for the 0% SPAIPT; (b) for 5%; (c) for 45%; (d) for 50%; (e) for 22.5 %; (f) for 25%; (g) for 27.5%. A SAI displays time interval histograms of the neural activity patterns (NAPs) for each of the tonotopic channels. In addition to these channel-by-channel patterns,
each panel depicts a summoned profile of the time interval histogram pooling over the tonotopic channel as well as a a spectral profile on the bottom, i.e., an excitation pattern, of the activity levels pooling over the time interval variation on the right side.

While the dominant peak of the time interval appeared naturally at 5 ms for the 0% SPAIPT (Fig. 6a), it shifted to the 2.5 ms point for the 50% SPAIPT (Fig. 6d). A deviation of 5% from each of these completely periodic stimuli impose some disturbance on the main peak as depicted in Fig. 6(b), and 6(c), respectively. Thus, the auditory model appeared to be able to describe the perceptual results qualitatively.

The 25% case, however, is puzzling. Although it’s main peak appeared at 5 ms point in Fig. 6(f), which is consistent with the experimental results depicted in Fig. 4 and 5, there was no peak around 2.5 ms corresponding to the pitch higher than the original by “one octave” which was likely to be perceived as a secondary option in the experiments. This does not suggest that there is some miss-implementation of the model. The model simply use time interval information observed in the neural activities to construct SAIs. Because the 25% SPAIPT does not contain any time interval corresponding to 50% of the cycle, it is natural that no peak appeared at 2.5 ms, i.e., 50% of the original 5 ms period.

Previous literatures have successfully revealed that the information of the temporal fine structures are necessary for our pitch perception, and representative models of pitch perception utilize this temporal information by taking some statistics of time intervals, such as, calculating autocorrelations, or adding delayed copies with event-driven triggers (implemented as in the strobed temporal integration). The results of such simulations have been usually depicted as a function of time intervals, or delays, which are simply physical measures. We need obviously at least another step to represent this physical property with certain neural codes.

The findings in Experiment 1 and 2 have provided a possibility that we could perceive a pitch for which no corresponding information of the temporal interval is found in the early auditory representation. This also suggest that it might be too simple to assume that pitch would be a direct correlate of the most prominent peak in the internal time interval histogram. The age-related shift of the absolute pitch would be another phenomenon which might raise a question about the relationship between the physical interval and the pitch percept.

5. EXPERIMENT 3: Age-related Shift of Absolute Pitch

Although some absolute pitch possessors anecdotally reported their sense of pitch judgment “shifted” when they were aged, the phenomena were not seriously attended partly because the possession of the absolute pitch (AP) are regarded as a quite exceptional abilities, partly because the number of systematic experimental findings had been quite small. As to the latter cause, tests...
Figure 5 – Stabilized Auditory Images for several examples of SPAIPTs.
recruiting a sufficient number of participants were done by using the internet technology and the results confirmed that the sense of AP shifts by aging. An interesting aspect of this age-related AP effect is that it is not a simple deterioration of judgments but is a systematic shift of judgments: AP possessors are likely to perceive pitches to be higher than they are young. As to the former aspect, we have a quite large numbers of AP possessors in Japan. The reason for such a cultural and/or regional difference has still been under investigation. It is, however, indicates that AP is not so exceptional ability as it had been believed, and that a series of systematic experiments can be done.

5.1. Purpose
Experiment 3 was done for two purposes: (1) to reconfirm that AP judgments shift affected by aging; (2) to investigate potential cues for the AP judgment. For the first purpose, a reasonable number of AP possessors whose ages ranged from 20s to 50s were recruited. For the second purpose, several kinds of experimental stimuli were used in the AP judgment tests in addition to real piano sounds.

5.2. Method
5.2.1. Participant
Seventy-two AP possessors participated in a series of experiments. They were recruited from the undergraduates and graduates of Faculty of Music, Kyoto City University of Arts, with a condition for self confidence of the AP possession. Their ages ranged from 19 to 58. The number of participants was 21, 15, 18, 18, for 20s (including 19 year old participants), 30s, 40s, and for 50s, respectively.

5.2.2. Stimulus
Seven kinds of test stimuli were used: (1) piano sounds; (2) pure tones (PT); (3) iterated rippled noises (IRN) (14); (4) missing fundamental harmonic complex tones of 5th, 6th, 7the, and 8th harmonics (MFHC_5678); (5) missing fundamental harmonic complex tones of 15th, 16th, 17the, and 18th harmonics (MFHC_15-18); (6) inharmonic complex tones of four components around 5th to 8th (IHC_5678); (9) inharmonic complex tones of four components around 15th to 18th (IHC_15-18). Except for the piano sounds, all the stimuli were digitally synthesized with a sampling frequency of 44.1 kHz and a quantization precision of 16 bits. Their schematic amplitude spectra are depicted in Fig. 7.

The piano sounds were sampled sounds of a real YAMAHA grand piano of 88 keys, i.e., ranging from 32 to 5373 Hz.
from A0 to C8. The fundamental frequencies of the other synthesized stimuli ranged from 65.4 Hz to 987.8 Hz, which corresponded to the range from C2 to B5 (4 octaves) with setting A4 at 440 Hz. To generate a series of the inharmonic complex tones, the frequency of each component was set at \((n_i + \alpha)F_0\) Hz, where \(n_i\) is the harmonic order, i.e., (5, 6, 7, 8) or (15, 16, 17, 18), \(\alpha = 0.25\), and \(F_0\) is the assumed fundamental frequency. Thus, the difference between each pair of the consecutive components was equal to the fundamental frequency. Although these stimuli were expediently called “inharmonic,” their component frequencies are, strictly speaking, harmonics of the fundamental of the quarter of the assumed one. The amplitude of each component was equalized for the MFHC and IHC series.

The IRNs were generated by repeating delay-and-add of white noise seed 256 times. The verse of the delay corresponded to the fundamental frequencies. The IRNs were low-pass filtered. The cut-off frequencies were set at the eight times of the assumed fundamental frequencies.

All of these synthesized stimuli had 250 ms duration with 50 ms linear rise/fall transitions. To minimize unexpected effects by the distortion product in the low frequency range, background low-pass noises were presented for the stimuli excepting the piano sounds and pure tones. Their cutoff frequencies were chosen from 250, 500, 1000, and 2000 Hz adaptively to the register of the assumed fundamental frequencies.

### 5.2.3. Procedure

All the participants were tested in two experimental sessions. In the first session, only the piano sounds were used. Each participant was required to judge the pitch name of a single piano sound from the twelve candidates with a semitone step, such as C, C#, D, D#, … and B by ignoring the pitch register. A session consisted of 88 trials. Each of the 88 pitch classes was presented once in a random order.

In the second session, the other synthesized stimuli were used. In each session, each stimulus combining the 6 types, 12 chroma, and 4 pitch registers was presented once in a random order. Thus, the total number of trials was 272. Each participant was required to judge the perceived pitch chroma, or nearby pitch chroma from 48 response choices dividing an octave. This fine-grained deviation was used because some AP possessors reported that slight deviation of pitch class from their “internal standard” especially for such artificial, experimental stimuli, and that they would help some difficulty if they were forced to choose one of the “established” pitch classes.

In both sessions, no feedback was provided.

### 5.3. Results

Although the participants responded all the 88 keys for the piano sounds, only the result for 48 keys ranging from C2 to B5 were used for the further analysis. For the synthesized stimuli except for the piano sounds, responses were quantized into 12 pitch chroma with a semitone step. The deviations from the “correct” pitch chroma were calculated.

Figure 8 depicts normalized histograms of these deviations for each stimulus type. The negative value corresponds to the case where a participant heard the pitch lower than it should be. The positive value corresponds to the case where he/she hear it higher than it should be. The difference in the gray scale of bars depicts the age group.

For the piano sounds, the proportion of the correct judgment decreased as the participants became older. This decrease of the correct judgment did not indicate simple degradation, as indicate by the asymmetrical increase of the positive side of the deviation by aging.

Similar patterns of the response distributions were observed for the stimulus type of PT, IRN, MFHC_5678. If one disregard the shift of judgment centers to “-1” point, the distribution pattern appeared similar to the stimulus type of IHC_5678.

On the other hand, the judgment distribution became flat for the stimulus type of MFHC_15-18 and IHC_15-18.

### 5.4. Discussion

For the piano sounds with which the participants should feel familiar, the proportion of the correct response decreased as the age increased. This age-related increase of misjudgments could
not be explained by a simple degradation in the precision of pitch judgment. The asymmetry of the misjudgment was observed. The misjudgment for the higher side increased as the age increased. Thus, the age effect of the AP judgment was reconfirmed.

A similar pattern was observed for either of the stimulus types, PT, MFC_5678, IRN, or IHC_5678. On the other hand, the manner of the response distributions indicated some ambiguity or instability of pitch for the MFC_15-18 and IHC_15-18. Therefore, the results suggest that the cue for the AP judgment might exist the lower harmonic region where one may expect a relatively better resolvability between components.

This implies that the age-related pitch shift might be caused by any modification of cues existing in rather low frequency region, which is unlikely to be explained by a common age-related problem in the high frequency region. In fact, it was reported that the correlation between the degree of AP shifts and the hearing level was not significant (15). It was also reported that the performance of frequency discrimination became worse in the aged group than in the young reference group even if they were matched in the audiometry (16). This suggests that the information of the fine temporal structure could be affected by aging.

It should, however, be noted that a degradation of the information in temporal fine structures could not explain a systematic deviation of pitch to a certain direction, i.e., to a higher direction as observed in Experiment 3. It would be worthy to mention that we have not found an appropriate way to represent physical time intervals in terms of any physiological code. Although the phase locked neural activities observable at the cochlea nerves can encode the frequency of vibration into temporal patterns, those temporal patterns should be coded into some other forms of neural codes to represent pitches. Meddis and O’Mard proposed a model which codes pitches with the chopper neurons and the coincidence detectors(17). The proposed model still codes differences in pitch in terms of patterns of activity distributions among the neurons with different chopping rates, and it remains still difficult to determine the perceived pitch simply by inspecting the output patterns. We would need further steps to predict the perceived pitch. Such approaches, however, are worthy to pursue.
6. **GENERAL DISCUSSION**

The results of Experiment 1 and 2 suggest that it would be too simplistic to relate the perceived pitch with the time interval having the most prominent peak activity. In addition, the results of Experiment 3 suggest that aging could affect the perceived pitch although the time interval should not be change. To explain the phenomena reported in the current paper, it seems necessary to assume further steps of the neural processing which could translate the activity patterns among the neurons responding different inter-activity intervals into activity distributions of a series of neurons each of which has the “best” chroma. It would be plausible to assume that a certain autonomic activity, such as chopper neurons, might be engaged in such a process. It would be likely for aging to affect the properties of such autonomic activities. Then, age-related modulation in such properties could induce the pitch shift.

**ACKNOWLEDGEMENTS**

This work was supported by JSPS KAKENHI Grant Number, 24243070 and 16H01734.

**REFERENCES**