CONTRIBUTIONS OF TEMPORAL AND PLACE CUES IN PITCH PERCEPTION OF ABSOLUTE PITCH POSSESSORS

PACS: 43.66.Hg

Fujisaki, Waka; Kashino, Makio

NTT Communication Science Laboratories, NTT Corporation,
3-1 Morinosato-Wakamiya, Atsugi, 243-0198 Japan
Tel: +81-46-240-4795; Fax: +81-46-240-4591
E-mail: fujisaki@avg.brl.ntt.co.jp (WF), mkashino@cslab.kecl.ntt.co.jp (MK)

ABSTRACT

Pitch perception is determined by both place and temporal cues. To explore if any differences exist in the use of these cues depending on absolute pitch capabilities and musical experiences, two pitch identification experiments with and without pitch references were conducted for subjects with different absolute pitch capabilities and musical experiences. Three types of stimuli were used to manipulate place and temporal cues separately: narrow-band noises, which provide a strong place cue but an ambiguous temporal cue; iterated rippled noises, which provide only temporal cue; and sinusoidal tones, which provide both. The results indicated that the temporal cue is important for chroma identification of absolute pitch possessors, and the place cue is important for height judgment of both absolute and non-absolute pitch possessors. Additionally, it is indicated that musical experts use both temporal and place cues more effectively when they identify both chroma and height of a musical pitch.

INTRODUCTION

Absolute pitch is the ability to identify or produce a pitch of a sound without hearing a reference sound. Generally, people tend to think that absolute pitch possessors are superior in all of their auditory and musical abilities. However, in the field of auditory and music psychology, researchers have focused on the etiology of absolute pitch or the merits and demerits of absolute pitch in music cognition, but have not closely studied the basic hearing mechanisms of absolute pitch possessors for general sound outside the context of music.

Studying the basic hearing abilities of absolute pitch possessors may bring insights not only into the mechanisms of absolute pitch itself, but also into problems such as individual differences and the plasticity of auditory functions. This possibility led us to conduct a series of experiments to examine whether absolute pitch capability has any influence on basic hearing abilities (Fujisaki & Kashino, 2002). The overall results showed that there were no significant differences in frequency, temporal and spatial resolutions among groups with different absolute pitch capability.

The next question then is where do the differences exist between absolute and non-absolute pitch possessors? Since absolute pitch is the ability to identify musical pitch, it is possible for differences from non-absolute pitch to exist in mechanisms directly related to musical pitch perception.

There are two major cues for pitch perception: "place cue" and "temporal cue" (Moore, 1997).

Sound waves reach the ear and cause the basilar membrane located in the cochlea of the inner ear to vibrate. The place where the basilar membrane vibrates varies according to the sound frequency. The peak of the vibration is near the end of the basilar membrane when frequencies are low, and is near the entrance (the oval window) when frequencies are high. There are many cells called "hair cells" on the basilar membrane, which transform the vibration into auditory nerve fiber impulses. More impulses are produced when the vibrations are large. Therefore, the amplitude spectrum of a signal reaching an ear can be coded by the excitation pattern of auditory nerve fibers along the basilar membrane. This information is called a "place cue", because it is based on the position on the basilar membrane. On the other hand, each auditory nerve fiber discharges impulses in synchrony with a certain phase of the basilar membrane vibrations when frequencies are below 4 to 5 kHz. Therefore, the periodicity of impulses that each auditory nerve fiber discharges contains the information of a frequency component of the input signal. This information is called a "temporal cue". In the case of a complex sound such as the sound of a musical instrument or a human voice, both types of cues work together to provide information that is used to determine pitch.

However, it is still unclear how absolute pitch possessors use temporal and place cues when they identify musical pitch, and whether any differences exist in the usage of these cues depending on the absolute pitch capabilities.

In this study, we conducted two experiments to answer these questions. In Experiment 1, we conducted relative pitch identification tasks in which a reference sound was presented before each target stimulus. In these tasks, place and temporal cues were separately manipulated. Then, the results across groups with different degrees of absolute pitch were compared. The reason we conducted relative pitch identification tasks rather than absolute pitch identification tasks in Experiment 1 was because non-absolute pitch possessors had difficulty performing in absolute pitch identification tasks. However, to examine how absolute pitch possessors use place and temporal cues when they identify musical pitch, we also needed to conduct absolute pitch identification tasks in which these cues were manipulated separately. This was done in Experiment 2, where we conducted absolute pitch identification tasks in which no reference sound was presented before each target stimuli. Absolute pitch and partial absolute pitch possessors participated in Experiment 2.

EXPERIMENT 1

Method

Twelve female musicians and six female non-musicians with normal hearing participated in this experiment. The musicians were subjects who had at least eight years of formal musical training before the age of twenty. The non-musicians were subjects who had less than two years of musical experience. Their age ranged from eighteen to thirty-two.

Before the main experiments, an absolute pitch test was conducted (note name identification task using pure tones in 3 octaves). Based on the results of the absolute pitch test, subjects were divided into three groups: absolute-pitch musicians, non-absolute-pitch musicians, and non-absolute-pitch non-musicians. Absolute-pitch musicians were musicians that scored more than 97% on the absolute pitch test, non-absolute-pitch musicians were musicians that scored under 25%, and non-absolute-pitch non-musicians were non-musicians that scored under 25%. Six of the musicians were diagnosed as absolute pitch, and the other six as non-absolute pitch. The six non-musicians were diagnosed as non-absolute-pitch non-musicians.

Three types of stimuli were used for pitch identification tasks: narrow-band noises, iterated rippled noises (the repeated broadband noises with a designated delay and gain), and pure tones. With narrow-band noises, the place cue dominates; the temporal cue is rather ambiguous. The bandwidth of narrow-band noises was 50 cents, and slope was 200 dB / octave. The iterated rippled noises had a limited bandwidth from 1000 to 3500 Hz. With iterated rippled noises, only the temporal cue is available as a cue for pitch. Iterated rippled noises, unlike narrow-band noises, contain regular periodic information. They can be designed by adding the original noise each time (Add-Original Network), or adding the latest repeated noise (Add-Same Network)(Yost, 1996). In this experiment, we used the latter approach. The repetition of iterated rippled noises was 16, and gain was 0.52. With pure tones, both place and temporal cues are useful in pitch perception. The signals were 250 ms in duration, shaped with 10-ms cosine ramps. The stimuli were presented binaurally via headphones (Sennheiser HDA 200), at a level of about 60 dB SPL.

The task was to identify the note name of a target sound by pressing the corresponding keys on a computer screen (keys were displayed in the same order as a musical keyboard). The task was almost the same as absolute pitch test, but a reference [middle C (261.6 Hz)] was presented before each target stimulus. The range of target stimulus was from C2 (65.4 Hz) to B6 (1975.5 Hz). The stimuli were narrow-band noises, iterated rippled noises, and pure tones.

Target stimuli in five octaves were presented in random order with the constraint that tones within the same octaves or tones having the same chroma were never presented successively. Absolute-pitch musicians, non-absolute-pitch musicians, and non-absolute-pitch non-musicians participated in this experiment. No feedback was returned. The order of stimuli conditions (narrow-band noises, iterated rippled noises, pure tones) was randomized for each subject. There were three trial sequences in each condition for each subject [one trial sequence contained 60 trials (60 notes in five octaves)], and the last two trial sequences were used for the analysis.

<u>Result and Discussion</u>

Figure 1 shows the results of this experiment. We calculated the percentage of correct responses in three ways: chroma correct (The note names of the presented target sound and the subject's response matched, ignoring absolute height); height correct (The note names of the target sound and the response were within ±5 semitones); and chroma & height correct (The note names of the target sounds and the responses matched exactly, including absolute height).

First, chroma correct scores were analyzed by a two-way analysis of variance. The interaction between groups (AP, NAP(M), NAP(NM)) and conditions (NBN, IRN, PT) was significant ($F(4, 30) = 13.35$, $p<.01$). Since the interaction between groups and conditions were significant, simple main effects for each factor were analyzed. First, simple main effects for groups were analyzed. In all groups, significant effect on stimuli was found (absolute-pitch musicians: $F(2,10)=22.25$, $p<.01$, non-absolute-pitch musicians: $F(2,10)=4.67$, $p<.05$; non-absolute-pitch non-musicians: $F(2,10)=23.24$, $p<.01$). From Tukey's HSD test on stimuli, in absolute-pitch musicians, there were significant differences between pure tones and iterated rippled noises ($p<.05$), and between iterated rippled noises and narrow-band noises ($p<.01$). There was no significant difference between pure tones and iterated rippled noises. In non-absolute-pitch musicians, there was a significant difference between pure tones and narrow-band noises ($p<.05$). In non-absolute-pitch non-musicians, there were significant differences between pure tones and iterated rippled noises ($p<.01$) and between pure tones and narrow-band noises.

Next, simple main effects for stimuli were analyzed. In all stimuli conditions, there were significant effects on all groups. From Tukey's HSD test on groups, there were significant differences between absolute-pitch musicians and non-absolute-pitch musicians (pure tones: $p<.01$; iterated rippled noises: $p<.01$, narrow band noises: $p<.05$), between absolute-pitch musicians and non-absolute-pitch non-musicians ($p<.01$ in all conditions), and between non-absolute-pitch musicians and non-absolute-pitch non-musicians ($p<.01$ in all conditions).

Narrow-band noises provide strong place cues but ambiguous temporal cues, and iterated rippled noises provide only temporal cues. Therefore, that the chroma score declined for narrow-band noises and did not decline for iterated rippled noises in the absolute pitch group suggests that temporal cues were very important for chroma identification for absolute pitch possessors. On the other hand, the chroma score declined for both narrow-band noises and iterated rippled noises for non-absolute-pitch non-musicians. This suggests that it was difficult for non-absolute-pitch non-musicians to identify chroma when only temporal cues were presented. They needed both temporal and place cues for chroma identification.

The results for non-absolute-pitch musicians were in between. The chroma score significantly declined for narrow-band noises compared to pure tones. However, there were no significant differences between narrow-band noises and iterated rippled noises nor iterated rippled noises and pure tones. This suggest that non-absolute-pitch musicians put more weight on temporal cues than non-absolute-pitch non-musicians and put less weight on temporal cues than absolute-pitch musicians when they identify chroma.

Then, height correct scores were analyzed by a two-way analysis of variance. The main effect of both groups (absolute-pitch musicians, non-absolute-pitch musicians, non-absolute-pitch non-musicians) and conditions (narrow-band noises, iterated rippled noises, pure tones) were significant (groups: $F(2,15)=6.15$, $p<.05$; conditions: $F(2,30)=56.03$, $p<.01$). The interaction between groups and conditions was not significant. From Tukey's HSD test on groups, there was a significant difference between absolute pitch musicians and non-absolute pitch non-musicians ($p<.01$). No significant differences were found between absolute-pitch musicians and non-absolute-pitch musicians, nor between non-absolute-pitch musicians and non-absolute-pitch non-musicians. From Tukey's HSD test on stimuli conditions, there were significant differences between narrow-band noises and iterated rippled noises ($p<.01$), and iterated rippled noises and pure tones ($p<.01$). No differences were found between pure tones and narrow-band noises.

Iterated rippled noises provide strong temporal cues, and narrow-band noises provide strong place cues but ambiguous temporal cues. Therefore, that the height score declined for iterated rippled noises and did not decline for narrow-band noises suggests that place cues play an important role in height identification.

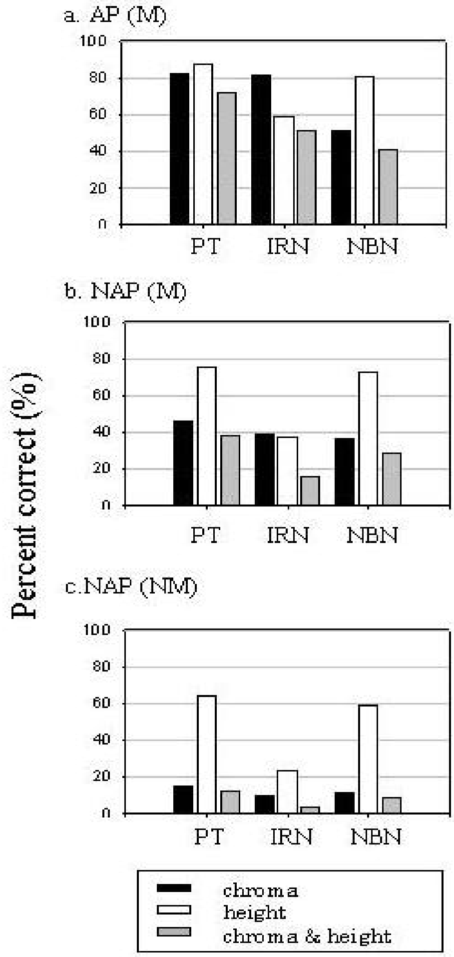

Figure 1
The percentages of correct responses with pure tones (PT), iterated rippled noises (IRN), and narrow-band noises (NBN) stimuli for absolute-pitch musicians [AP(M)], non-absolute-pitch musicians [NAP (M)] and non-absolute-pitch non-musicians [NAP (NM)] listeners are shown. The percentages of correct responses were calculated in three ways: chroma correct, height correct, and chroma and height correct.

EXPERIMENT 2

Method

Fourteen female musicians with normal hearing participated in Experiment 2. Like the musicians in Experiment 1, they had at least eight years of formal musical training before the age of twenty. Based on the results of the absolute pitch test, subjects were divided into two groups: absolute-pitch musicians, and partial-absolute-pitch musicians. Absolute-pitch subjects were musicians that scored more than 97% on the absolute pitch test, and partial-absolute-pitch subjects were those that scored between 65 % and 85%. Seven were diagnosed as absolute pitch and the other seven as partial-absolute pitch. Four of the seven absolute-pitch subjects also participated in Experiment 1. The same three types of stimuli were used for pitch identification tasks.

The task was to identify the note name of a target sound by pressing the corresponding keys on a computer screen. The task was almost same as in Experiment 1, but no reference was presented before each target stimulus. As in Experiment 1, the range of target stimulus was from C2 (65.4 Hz) to B6(1975.5 Hz).

Target stimuli in five octaves were presented in random order with the constraint that tones within the same octaves or tones having the same chroma were never presented successively. No feedback was returned. The order of stimuli conditions (narrow-band noises, iterated rippled noises, pure tones) was randomized for each subject. There were three trial sequences in each condition for each subject [one trial sequence contains 60 trials (60 notes in five octaves)], and the last two trial sequences were used for the analysis.

Result and Discussion

Figure 2 shows the results of this experiment. As in Experiment 1, we calculated the percentage of correct responses in three ways: chroma correct, height correct, and chroma & height correct.

First, chroma correct scores were analyzed by a two-way analysis of variance. The main effect of both groups (absolute pitch, partial-absolute pitch) and conditions (narrow-band noises, iterated rippled noises, pure tones) were significant (groups: $F(1,13)=24.56$, $p<.01$; conditions: $F(2,26)=17.49$, $p<.01$). The interaction between groups and conditions was not significant. Since the effect of groups was significant, it was indicated that absolute-pitch possessors scored higher than partial-absolute-pitch possessors in chroma identification. From Tukey's HSD test on stimuli conditions, there were significant differences between narrow-band noises and between iterated rippled noises ($p<.01$), and iterated rippled noise and pure tones($p<.01$). No differences were found between pure tones and narrow-band noises.

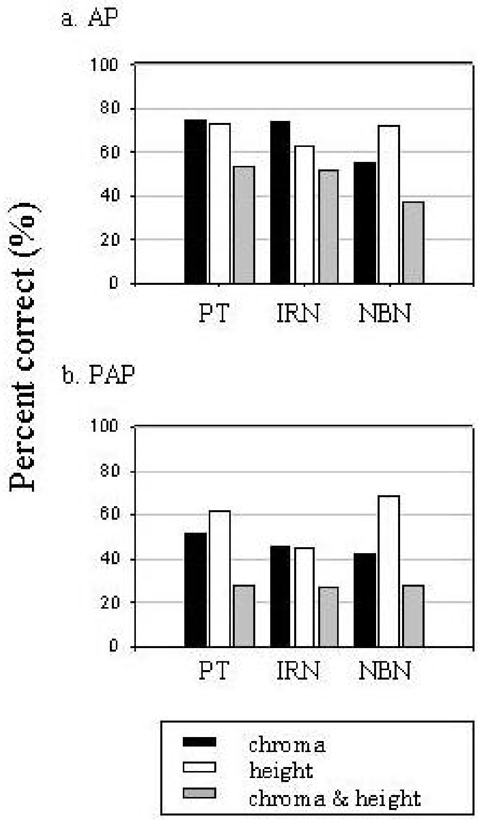

Figure 2. Results of Experiment 2
The percentages of correct responses with PT, IRN, and NBN stimuli for AP and PAP (partial-absolute pitch) listeners are shown. Same as Exp.1, the percentages of correct responses were calculated in three ways: chroma correct, height correct, and chroma and height correct.

Iterated rippled noises provide strong temporal cues, and narrow-band noises provide strong place cues but ambiguous temporal cues. Therefore, the result that the chroma score declined for narrow-band noises and did not decline for iterated rippled noises suggests that temporal cues play an important role in chroma identification for absolute-pitch possessors.

Next, height correct scores were analyzed by a two-way analysis of variance. The main effect of the conditions (narrow band noises, iterated rippled noises, pure tones) was significant ($F(2,26)=10.81$, $p<.01$). However, neither the main effect of the groups nor the interaction between groups and conditions was significant. Since there was no significant difference between the groups, it appears that unlike the chroma identification ability, the height identification ability of absolute-pitch possessors and partial-absolute-pitch possessors were not significantly different.

From Tukey's HSD test on stimuli, height score for iterated rippled noises was significantly lower than the scores for pure tones and narrow-band noises. Iterated rippled noises provide strong temporal cues, and narrow-band noises provide strong place cues but ambiguous temporal cues. Therefore, the result that height score declined for iterated rippled noises and did not decline for narrow-band noises suggests that place cues play an important role for height identification.

GENERAL DISCUSSION

This study indicated that the temporal cue were important for chroma identification of absolute pitch possessors, and the place cue was important for height judgment of both absolute and non-absolute pitch possessors. Additionally, it was indicated that musical experts use both temporal and place cues more effectively when they identify both the chroma and height of musical pitch.

One possible interpretation of these findings is that musical experiences may increase the usage of both temporal and place cues in musical pitch identification, but when the weight on the temporal cues is more emphasized, the ability of absolute pitch may be acquired. When a listener perceives chroma, or musical intervals, such as an octave or a perfect fifth, it is quite natural to assume that the temporal cue is particularly important (Ohgushi, 1983). Notes with an interval such as an octave (the frequency ratio is 1:2) or a perfect fifth (the frequency ratio is 2:3) are similar from the viewpoint of the periodicity of auditory nerve firing, but it is not necessary that these frequency ratios become especially significant points from the viewpoint of the position on the basilar membrane. Also, the upper limit of chroma perception of absolute pitch possessors is about 5 kHz (Bachem, 1948) and it almost agrees with the limit of phase locking (Rose, Brugge, Anderson & Hind, 1968). This may reflect the importance of temporal cues in chroma perception.

A hypothesis can now be made. When children paid more attention to temporal cues than place cues and learned the connection between the temporal cues and its note name, they may acquire absolute pitch. To evaluate this hypothesis, further studies from various viewpoints that take into account the development of neural systems and the plasticity of auditory functions are needed.

In this study, we found that absolute pitch possessors can utilize temporal cues more effectively when they identify musical pitch. The next question then is why they can use temporal cues more effectively? Can absolute pitch possessors detect temporal cues better than non-absolute pitch possessors? Or, are there no differences in the detection ability of temporal cues? Instead do differences exist in the process where temporal cues are used for pitch name identifications? To clarify these points, further studies that examine the discrimination and detection abilities of place and temporal cues are needed.

REFERENCES

Bachem, A. (1948). Chroma fixation at the ends of the musical frequency scale. *Journal of the Acoustical Society of America, **20***, 704-705.

Fujisaki, W. & Kashino, M. (2002) The basic hearing abilities of absolute pitch possessors. *Acoustical Science & Technology*, **23**(2), 77-83.

Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing* (4th ed. ed.). San Diego: Academic Press.

Ohgushi, K. (1983). The origin of tonality and a possible explanation of the octave enlargement phenomenon. *Journal of the Acoustical Society of America, **73** (5)*, 1694-1701.

Rose, J. E., Brugge, J. F., Anderson, D. J. & Hind. J. E. (1968). Patterns of activity in single auditory nerve fibers of the squirrel monkey. In A. V. S. de Reuck and J. Knight (Ed.), *Hearing Mechanisms in Vertebrates* . Churchill, London.

Yost, W. A. (1996). Pitch of iterated rippled noise. *Journal of the Acoustical Society of America, **100***, 511-518.