Basic hearing abilities and characteristics of musical pitch perception in absolute pitch possessors

Waka Fujisaki & Makio Kashino

NTT Communication Science Laboratories, NTT Corporation, Japan
fujisaki@avg.brl.ntt.co.jp, kashino@avg.brl.ntt.co.jp

Abstract

Two studies examined the basic hearing abilities and characteristics of musical pitch perception in absolute pitch possessors. First, we investigated whether absolute pitch capabilities have any influence on basic hearing abilities with respect to frequency, temporal, and spatial resolution. The overall results showed that there were no significant differences among groups with different absolute pitch capability in terms of resolutions of the auditory system. Second, we explored whether differences from non-absolute pitch exist in the mechanisms directly related to musical pitch perception. Pitch perception is determined by both place and temporal cues. The results indicated that absolute pitch possessors could utilize temporal cues more effectively when they identify musical chroma. With regard to the judgment of height, it was found that the place cues play an important role for both absolute and non-absolute pitch possessors.

1. Introduction

Absolute pitch is the ability to identify or produce a pitch of a sound without hearing a reference sound [1], [2]. Absolute pitch can be said to be one example of the diversity of auditory perception. There are many different kinds and levels of absolute pitch; some people can identify one or two musical pitches, some all white keys of a piano keyboard, some all black and white keys, and some not only piano sounds but also sounds such as an ambulance siren, birds singing, or the sound of a struck.

Absolute pitch can also be considered on example of the plasticity of auditory functions because absolute pitch is usually acquired before a critical period at around the age of 3-6. Therefore, studying absolute pitch may provide insights not only into the mechanisms of absolute pitch itself, but also into problems, such as the diversity and plasticity of auditory functions. This led us to conduct a series of studies to examine the basic hearing abilities and characteristics of musical pitch perception among the people with different absolute pitch capabilities.

Generally, absolute pitch possessors tend to be thought of having ‘good ears’ in all of their auditory abilities. However, it has not been systematically investigated whether absolute pitch capabilities have any influence on basic hearing abilities for general sound outside the context of music. In the first study, we conducted four experiments to examine whether any differences exist in the resolutions of the auditory system among people with different absolute pitch capabilities [3]. We measured fundamental elements of auditory abilities, including frequency resolution (the ability to detect small frequency differences in sounds), temporal resolution (the ability to detect fast temporal changes in sounds) and spatial resolution (the ability to detect small differences in the spatial localization of sounds).

In the second study, we focused on the mechanisms directly related to musical pitch perception [4], [5]. Pitch perception is determined by both place and temporal cues [6]. When the sound wave reaches the ear, it causes the basilar membrane located in the cochlea to vibrate. On the basilar membrane, inner hair cells transform the vibration into impulses of auditory nerve fibers. When the vibrations are large, more impulses are produced. Therefore, the amplitude spectrum of the signal can be coded by the excitation pattern of auditory nerve fibers along the basilar membrane. This information is referred to as ‘place cues’, 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 lower than 4 to 5 kHz. The periodicity of auditory nerve firing also contains the information about the frequency component of the input signal. This information is referred to as ‘temporal cues’.

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. To explore if any differences exist in the usage of these acoustic cues for musical pitch identification, we conducted two experiments with and without pitch references for subjects with different absolute pitch capabilities.

2. Basic hearing abilities of AP

Before the main experiments, an absolute pitch test was conducted. Pure tones in three octaves were presented in a random order with the constraint that tones within the same octaves or tones having the same chroma were never presented successively. Before
conducting the absolute pitch test, it was confirmed that all subjects knew the correspondence between keys of the musical keyboard and musical pitch names. Subjects were asked to identify and choose the musical pitch by clicking buttons on a computer screen. When subjects could not identify the pitch name of the target sound, they were instructed to choose the key that they thought was closest to the target sound.

Based on the results of the absolute pitch test, subjects were divided into four groups: absolute pitch (AP) musicians, partial AP (PAP) musicians, non-AP (NAP) musicians, and non-AP non-musicians (NM). AP subjects were musicians that scored more than 97% on the absolute pitch test, PAP subjects were musicians that scored between 65-85%, NAP subjects were musicians that scored under 25%, and NM subjects were non-musicians that scored under 25%.

We measured the following four attributions, which are used in measuring the frequency resolution, temporal resolution and spatial resolution:

1. Frequency Discrimination Thresholds: Two sinusoidal tones with different frequencies centered on a test frequency \( f_t \) were presented to measure the frequency discrimination thresholds. \( f_t \) was 1000 Hz or 987.76 Hz (The latter corresponds to B5 in the equal tempered scale). The task of subjects was to indicate the higher interval from two successive intervals of tones.

2. Tone in Notched-Noise Detection Thresholds: The thresholds of a signal masked by a noise with a bandstop (notch) were measured by changing the notch width. Two successive intervals of notched noise were presented in which either the first or the second interval contained a signal, and subjects judged which interval contained the signal.

3. Gap Detection Thresholds: Two successive intervals of white noise were presented in which either the first or the second interval was interrupted to produce the gap. The subjects’ task was to indicate which interval contained a gap.

4. Interaural Time Differences Discrimination Thresholds: Two successive intervals of pure tones were presented where either the first or second interval had larger ITD (A positive value of ITD indicates that the right ear signal leads). The task was to indicate which intervals were perceived further right.

Results of these experiments are shown in figure 1. The overall results show that there were no significant differences in frequency, temporal, and spatial resolutions among groups with different absolute pitch capability. Although in the frequency discrimination experiment, the effect of musical experience was found when test frequency \( f_t \) was in-tune to the equal tempered scale.

Absolute pitch possessors tend to be thought of having ‘good ears’, but our findings show that absolute pitch possessors do not have particularly ‘good ears’ in terms of frequency, temporal, and spatial resolutions at least for the tasks in the present study.

Figure 1: Frequency discrimination, tone in notched noise detection, gap detection and ITD discrimination thresholds obtained with AP, PAP, NAP and NM subjects.
3. Characteristics of musical pitch identification in AP

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. Thus, we examined what information absolute-pitch possessors use when they identify musical pitch and how the usage is different from non-absolute-pitch possessors.

Three types of stimuli were used for pitch identification tasks. Narrow band noises (NBN) were used as the stimuli in which pitch perception is mainly produced by place cues, and temporal cues are rather less salient. The bandwidth of narrow band noises was 50 cents, and slope was 200 dB/octave. Repeated broadband noises with a designated delay and gain and limited bandwidth from 1000 to 3500 Hz (Iterated rippled noises (IRN)) were used as the stimuli in which pitch perception is mainly produced by temporal cues, and place cues are rather less salient. IRN, unlike NBN, contains regular periodic information. IRN with 1000/d Hz pitch can be designed by repeating the process of adding noise with d ms delay (100-Hz pitch is perceived with 10-ms delay). IRN can be designed by adding the original noise each time (Add-Original Network) or adding the latest repeated noise (Add-Same Network) [7]. In this experiment, we used the latter approach. This noise adding process was repeated 16 times, and the gain of the repeated noise was about 0.52 of the latest repeated noise. With pure tones (PT), pitch perception is produced by both place and temporal cues.

The task was to identify the note name of a target sound by pressing the corresponding keys on a computer screen. In Expt. 1, middle C (261.6 Hz) was presented as the reference sound before each target stimulus. In Expt. 2, no reference sound was presented. AP, NAP and NM subjects participated in Expt. 1, and AP and PAP subjects participated in Expt.2.

The numbers of correct responses were calculated in two ways. Chroma correct responses were defined as the responses where the note names of the presented target sounds and the subjects’ responses matched, ignoring absolute height (The chance level was about 8.3 %). Height correct responses were defined as the responses where the note names of the target sounds and the responses were within ±5 semitones (The chance level was about 18.3 %).

Chroma and height correct scores in Expt. 1 are shown in figure 2. With regard to the identification of chroma, the interaction between groups and stimulus conditions was significant. In the AP group, there were significant differences between pure tones and narrow band noises (p<.01), and between iterated rippled noises and narrow band noises (p<.01). There was no significant difference between pure tones and iterated rippled noises. In the NAP group, there was a significant difference between pure tones and narrow band noises (p<.05). In the NM group, there were significant differences between pure tones and iterated rippled noises (p<.01) and pure tones and narrow band noises (p<.01).

Narrow band noises provide strong place cues but less salient temporal cues, and iterated rippled noises provide strong temporal cues and less salient place cues. Therefore, the results in which chroma score declined for narrow band noises and did not decline for iterated rippled noises in the AP group suggest that temporal cues are very important for chroma identification for absolute pitch possessors.

On the other hand, the chroma score declined in both narrow band noises and iterated rippled noises conditions in the NM group. This suggests that it was difficult for NM subjects to identify chroma when only temporal cues were presented.

The results for NAP musicians were in between. The chroma scores 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 NAP musicians put more weight on temporal cues than NM and put less weight on temporal cues than AP when they identify chroma.

With regard to the judgment of height, there was a significant difference between AP and NM (p<.01); however, no significant differences were found between AP and NAP, nor NAP and NM. The interaction between groups and conditions was not significant. 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 but less salient place cues, and narrow band noises provide strong place cues but less salient temporal cues. Therefore, the results that height score declined for iterated rippled noises condition and did not decline in narrow band noises condition suggest that place cues play an important role for height identification.

Figure 3 shows chroma and height correct scores in Expt. 2 where AP and PAP subjects participated in pitch identification experiment where no reference sound was presented. The overall tendencies were the same as the results of Expt. 1, where the reference sound was presented before each target stimulus.

4. Discussion

Our findings suggest that absolute pitch possessors may not have particularly ‘good ears’ in terms of resolutions. However, the differences from non-absolute pitch may originate in the relative weights on different acoustical cues for pitch identifications. It is
indicated that temporal cues are important for chroma identification by absolute pitch possessors, and that place cues are important for height judgment by both absolute and non-absolute pitch possessors.

One possible interpretation of these findings is that musical experience 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 [8]. Musical 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 [9], which almost agrees with the limit of phase locking. This may reflect the importance of temporal cues in chroma perception.

In this study, we found that absolute pitch possessors can utilize temporal cues more effectively when they identify musical chroma. This then prompts a few questions. Why can absolute pitch possessors utilize temporal cues more effectively? Do the differences from non-absolute pitch exist in the detection ability of temporal cues? Or do the differences exist not in the detection ability but in the process where temporal cues are used for pitch name identifications? Clarifying these points requires further studies to examine the discrimination and detection abilities of place and temporal cues.

5. References