Implications for pitch mechanisms of perceptual learning of fundamental frequency discrimination: Effects of spectral region and phase

Hiromitsu Miyazono1,* and Brian C. J. Moore2

1Department of Administration, Prefectural University of Kumamoto, 3–1–100 Tsukide, Higashi-ku, Kumamoto, 862–8502 Japan
2Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 2EB, United Kingdom

(Received 14 January 2013, Accepted for publication 19 June 2013)

Abstract: Perceptual learning was used to examine mechanisms of pitch perception. Thresholds (F0DLs) were measured for discrimination of the fundamental frequency (F0) of complex tones with a nominal F0 of 100 Hz and cosine-phase or random-phase harmonics. Tones were bandpass filtered and presented in threshold equalizing noise. A group trained using stimuli with the filter centered on LOW harmonics (1–5) showed a large training effect, with transfer to stimuli with MID harmonics (11–15) or MID-HIGH harmonics (14–18), but no transfer to stimuli with HIGH harmonics (28–32). A group trained with MID or MID-HIGH stimuli showed a large training effect, with transfer to the LOW stimuli and no transfer to the HIGH stimuli. A group trained with HIGH stimuli showed no training effect for any stimuli. The results suggest that similar mechanisms were used for F0 discrimination of the LOW, MID, and MID-HIGH stimuli, and that a different mechanism was used for the HIGH stimuli. It is proposed that the LOW, MID and MID-HIGH stimuli were discriminated using temporal fine structure (TFS) information, in the former case TFS information about individual resolved harmonics, and in the latter two cases TFS information about the periodicity of the waveform evoked by interfering harmonics.

Keywords: Perceptual learning, Pitch, Temporal fine structure, Resolvability, Training effect

PACS number: 43.66.Hg [doi:10.1250/ast.34.404]

1. INTRODUCTION

“Perceptual learning” is the phenomenon that participants tend to improve with training on many perceptual tasks, including frequency discrimination [1]. The design of perceptual learning studies usually involves a pre-training session in which discrimination performance is assessed for stimuli in two or more categories, where it is hypothesized that different perceptual mechanisms might underlie the discrimination of the stimuli in, say, categories A and B. Then participants are split into two groups. Group 1 is trained using a stimulus from category A, while group 2 is trained using a stimulus from category B. Then, in a post-training session, performance is assessed once more for all stimuli in both categories. If there are different perceptual mechanisms for processing the A and B stimuli, then group 1 should show greater improvements (relative to performance in the pre-training session) for stimuli in category A than for stimuli in category B, while group 2 should show greater improvements for stimuli in category B than for stimuli in category A. It should be noted that this type of design is based on the assumption that the perceptual mechanisms involved are independent and that a given type of stimulus will be processed using either one mechanism or the other. The situation may be different when two mechanisms operate in series [2].

One study with the above design was described by Demany and Semal [3]. In their experiment, participants were trained to discriminate the fundamental frequency (F0) of complex tones containing harmonics 3–7. Then, frequency discrimination was assessed for a pure tone whose frequency was either equal to the (missing) F0 of the complex with which they were trained or remote from it. Similarity in pitch of the complex used for training and the sinusoid used in subsequent testing did not seem to promote the transfer of learning. Demany and Semal
concluded that pitch discrimination learning is at least partly timbre-specific or spectrum-specific.

Some researchers have proposed that there are two mechanisms for the perception of the pitch of complex tones, one for low harmonics, which are usually resolved in the peripheral auditory system, and one for high harmonics, which are unresolved [4,5]. It has been suggested that a harmonic can be resolved, in the sense that it can be “heard out” from a complex tone with 75% accuracy, when it is separated from neighboring harmonics by 1.25ERB_N or more [6,7], where ERB_N is the equivalent rectangular bandwidth of the auditory filter at medium sound levels for young listeners with normal hearing [8,9]. This means that, depending on the F0, the lowest 5–8 harmonics are resolved, whereas higher harmonics are not resolved [10]. Other definitions of resolvability lead to similar estimates [11], although Bernstein and Oxenham [12] proposed that harmonics up to about the 10th could be resolved for F0s of 100 and 200 Hz.

For tones with low harmonics, the pitch may be derived from temporal fine structure (TFS) information (phase locking) or place information about the frequencies of individual harmonics [13], while for tones with very high harmonics the pitch may be derived from the temporal envelope of the waveform evoked on the basilar membrane [14–16]. However, some researchers have proposed that pitch may be extracted in a similar way for high and low harmonics [9,17]. If there are different pitch mechanisms for low and high harmonics, then training on F0 discrimination of tones with high harmonics should lead to improvements in F0 discrimination mainly for tones with high harmonics; the learning will not transfer or generalize to tones with low harmonics because of the different mechanisms involved. Similarly, training on F0 discrimination of tones with low harmonics should not transfer to tones with high harmonics. However, if there is a single pitch mechanism for low and high harmonics, then training using tones with high harmonics should transfer (i.e., lead to better F0 discrimination) to tones with low harmonics and vice versa.

Grimald et al. [18] conducted a study based on this idea. One group of participants was trained on F0 discrimination using complex tones with F0 = 88 Hz, filtered into the frequency range 1,375–1,875 Hz. This tone contained only high unresolved harmonics. A second group was trained using complex tones with F0 = 250 Hz, filtered into the same frequency range. This tone contained a few resolved harmonics. The components were added starting in sine phase (0°), which leads to a waveform with a high peak factor on the basilar membrane when the components are unresolved. It was found that, for both groups, learning occurred for the trained tones, and transferred somewhat to the other complex tones. However, for the group trained using a tone with only unresolved harmonics, the transfer of learning was greater for tones with unresolved than for tones with resolved harmonics. In contrast, for the group trained using a tone with some resolved harmonics, the transfer of learning was greater for tones with resolved than for tones with unresolved harmonics. They argued that their results supported the idea that there are two pitch mechanisms, one for resolved and the other for unresolved harmonics. Also, they argued that the general transfer of learning that was observed might reflect procedural learning or might reflect “some generic underlying F0-discrimination mechanism that does not strictly depend on the frequency region and the F0 being tested.”

Miyazono and Moore [19] measured thresholds (F0DLs) for discrimination of the F0 of a group of harmonics (Group B) embedded in harmonics with a fixed F0, equal to the mean F0 for group B. They found a large training effect for tones with high harmonics in Group B when the harmonics were added in cosine phase. Miyazono et al. [20] showed that this effect was due to use of a cue related to pitch pulse asynchrony (PPA), the degree to which the envelope peaks of waveforms on the basilar membrane are synchronised across different places (different frequency channels). When PPA cues were disrupted by introducing a temporal offset between the envelope peaks of the harmonics in Group B and the remaining harmonics, F0DLs increased markedly. Miyazono et al. [20] examined perceptual learning using a training stimulus with cosine-phase harmonics, F0 = 50 Hz, and high harmonics in Group B, under conditions where PPA was not useful. Learning occurred, and it transferred to other cosine-phase tones, but not to random-phase tones. A similar experiment with F0 = 100 Hz showed a learning effect that transferred to a cosine-phase tone with mainly high unresolved harmonics, but not to cosine-phase tones with low harmonics, and not to random-phase tones. The learning found by Miyazono et al. [20] appeared to be specific to tones for which F0 discrimination was based on distinct peaks in the temporal envelope.

So far, we have focused on the possibility that there are two pitch mechanisms. However, one can argue that there might be three mechanisms. For tones with intermediate harmonics, with harmonic numbers in the range 9 to 14, the harmonics are probably unresolved [6,7,10,11,21], but the pitch may be extracted from TFS information relating to the time interval between prominent peaks in the waveform evoked on the basilar membrane by the interference of harmonics [15,16,22–24]. Specifically, the pitch is assumed to be related to the time intervals between peaks in the TFS close to adjacent envelope maxima. For example, if the F0 is 100 Hz, and the tone is bandpass filtered to contain harmonics around 1,000 Hz, the time intervals are close...
to 9, 10 and 11 ms. A pitch corresponding to 100 Hz (1/10 ms) is usually heard, but pitches corresponding to 111 Hz (1/9 ms) and 90.9 Hz (1/11 ms) may also be heard [15]. When the harmonic numbers are too high (above the 14th), it appears that this form of TFS information cannot be used, and only information related to the temporal envelope can be used to extract pitch [23]. Pitch processing of stimuli whose harmonics are (just) unresolved, but not above the 14th, may reflect a third type of pitch mechanism, or it may reflect similar processing to that involved for low resolved harmonics.

In the present study, we set out to test these possibilities. We studied perceptual learning for F0 discrimination using complex tones that were bandpass filtered so as to contain low resolved harmonics (stimulus LOW), high unresolved harmonics (stimulus HIGH), and intermediate harmonics (stimulus MID). All stimuli were presented in a background of threshold equalizing noise (TEN) [25] to mask combination tones and to limit the audibility of components falling outside the passband. In Experiment I, the filters were chosen to have relatively shallow slopes of 30 dB/oct. The shallow slopes were used so that, when the harmonics were unresolved, changes in F0 would result in minimal changes in the excitation pattern [23,26]. Also, the use of shallow slopes meant that there were no “edge” harmonics (harmonics with no adjacent harmonics above or below them), avoiding the possibility that edge harmonics might be unusually well resolved [6].

2. EXPERIMENT I: TRAINING EFFECTS WITH HARMONICS 11–15 IN THE MID STIMULUS

2.1. Stimuli

In each trial, participants were presented with two successive tones, each with rise/fall times of 20 ms and a steady duration of 500 ms. The complex tones had an overall level of 65 dB SPL. The nominal F0 was 100 Hz. Three fixed spectral envelopes were used for the test stimuli, each with a flat bandpass region and slopes of 30 dB/octave. The flat region was 400-Hz wide, so 5 harmonics fell within the passband. For cases LOW, MID and HIGH, the passbands extended from 100 to 500, 1,100 to 1,500, and 2,800 to 3,200 Hz, respectively. All components were added with cosine starting phase. The TEN spectrum ranged from 100 to 8,000 Hz. The TEN level at 1 kHz, expressed as dB/ERBN [9], was set 20 dB below the level of the each component within the passband. Figure 1 shows a schematic spectrum for the MID stimulus.

2.2. Participants

Participants were 15 university students with absolute thresholds better than 10 dB HL over the range of audiometric frequencies from 250 to 8000 Hz. Their ages ranged from 21 to 23 years. They were not musically trained. Participants were divided into three groups of five participants, designated LOW, MID and HIGH, according to the stimuli used during training (see below).

2.3. Procedure

F0DLs were measured using a 2-interval, 2-alternative forced-choice procedure with feedback. Participants were asked to indicate which of the two tones in a trial they perceived to have the higher pitch. A three-down, one-up adaptive procedure was used to estimate the 79% correct point on the psychometric function. The mean F0 of the two tones to be discriminated was 100 Hz. The difference in F0 between the two tones to be discriminated, $\Delta F$, was changed by a factor of $2^{0.5}$ until four turnpoints had occurred, and was changed by a factor of $2^{0.25}$ thereafter. Twelve turnpoints were obtained and the geometric mean of the values of $\Delta F$ at the last eight turnpoints was used to estimate the F0DL. Each participant was tested on 10 days, 2 for measurement of pre-training thresholds, 6 for training, and 2 for measurement of post-training thresholds. There were 2 days between each testing day. In the pre-training and post-training sessions, all 15 participants were tested using all three conditions (LOW, MID, and HIGH). In each of these sessions, the F0DL was estimated at least 3 times for each condition. The training was conducted on successive days from Monday to Friday, but there were breaks in training on Saturday and Sunday. On each training day there were two sessions, separated by about 20 minutes. Three threshold estimates were obtained in each session, giving six estimates per day and 36 estimates in total for the training. Following training, participants were tested again using the same conditions as for pre-testing. Participants were divided into three groups of five participants for the training sessions. Table 1 lists the harmonics present in the passband during pre/post-testing and training for each group.
2.4. Results

Figure 2 shows the results obtained for the pre-training session (Pre), the training sessions, and the post-training session (Post), for each group. Thin curves show geometric mean F0DLs for the individual participants, and thick curves show the geometric mean across participants. The F0DLs are expressed as relative values in % ($\frac{100}{C3} \frac{F}{C1} F = F_0$).

Performance improved across days for groups LOW and MID, but not for group HIGH. An analysis of variance with group membership (LOW, MID, HIGH) as a between-subject factor and session as a within-subject factor showed a significant effect of session for groups LOW [$F(11,44) = 2.92$, $p < 0.01$], and MID [$F(11,44) = 3.11$, $p < 0.001$], but not for group HIGH [$F(11,44) = 0.23$, $p > 0.05$].

Figure 3 shows the overall training effect for each group and each stimulus type, expressed as the mean F0DL for the pre-training session divided by the mean F0DL for the post-training session. The three sets of bars represent the three stimulus cases, and the three bars within each set represent the three groups. Group LOW showed a large training effect for the LOW stimuli, with strong transfer to the MID stimuli, but no transfer to the HIGH stimuli. Group MID showed a large training effect for the MID stimuli, with strong transfer to the LOW stimuli and no transfer to the HIGH stimuli. Group HIGH showed no training effect for any stimuli.

An analysis of variance was conducted on the ratios shown in Fig. 3, with group membership as a between-subject factor and stimulus type as a within-subject factor. This and all subsequent analyses were based on the logarithms of the ratios, to make the variance more uniform across conditions. With this transform, Mauchley’s test for sphericity did not indicate any need to adjust the degrees of freedom. There was a significant effect of group; $F(2,12) = 32.8$, $p < 0.001$. Post hoc tests (bonferroni corrected here and in later analyses) showed that the...
A training effect was smaller for group HIGH than for groups LOW or MID (both \( p < 0.001 \)), but the effect did not differ for groups LOW and MID. The effect of stimulus type was significant; \( F(2, 24) = 34.2, p < 0.001 \). The training effect was smaller for stimulus HIGH than for stimuli LOW or MID (both \( p < 0.001 \)), but the effect did not differ for stimuli LOW and MID. The interaction of group and stimulus type was significant; \( F(4, 24) = 12.46, p < 0.001 \). For groups LOW and MID, the training effect was greater for stimuli LOW and MID than for stimulus HIGH (both \( p < 0.01 \)). For group HIGH, the training effect did not differ across the three stimulus types.

2.5. Discussion

The fact that there was no training effect for Group HIGH, while there was for groups LOW and MID, suggests that the mechanism underlying F0 discrimination was different for the HIGH stimuli than for the LOW or MID stimuli. It is possible that learning for the HIGH stimuli occurred so quickly that the learning was complete by the end of the pre-training session. However, this seems unlikely, given that the stimulus for case HIGH had a very weak pitch, making it harder to discriminate changes in F0. Hence, the most likely explanation for the lack of a learning effect for case HIGH stimuli, and the lack of transfer of learning for group HIGH to the LOW and MID stimuli, is that the mechanism underlying F0 discrimination was different for the HIGH stimuli than for the LOW or MID stimuli.

In Experiment I, the passband for stimulus MID contained harmonics 11 to 15. The harmonics in this stimulus were largely unresolved, and F0 discrimination was probably based on TFS information derived from unresolved harmonics. Hence, the similarity of the training effect for cases LOW and MID, and the transfer of learning between these two cases, supports the idea that F0 discrimination was based on a common mechanism, perhaps using TFS information (from resolved harmonics for stimulus LOW and unresolved harmonics for stimulus MID). However, harmonics 7, 8, 9 and 10, which fell on the slope below the passband, would have been above the masked threshold in the TEN. Bernstein and Oxenham [12] suggested that harmonics up to the 10th might be resolved, although there are some problems with the method used by them [27,28]. It is possible, therefore, that the lowest audible harmonics in stimulus MID were resolved to some extent. This could account for the similarity of the results for the LOW and MID stimuli, and the transfer of learning for these two types of stimuli. To assess this possibility, in Experiment II we measured learning and transfer effects using new MID stimuli, denoted MID-HIGH, filtered so that the passband contained harmonics 14–18. In addition the lower spectral slope was made steeper, being 60 dB/oct rather than 30 dB/oct. This meant that fewer harmonics falling on the lower slope were above the masked threshold in the TEN. The lowest audible harmonic in the MID-HIGH stimuli was the 11th. This would not have been resolved, but the frequency region of the lowest audible harmonics might have been low enough for TFS information to be used.

3. EXPERIMENT II: TRAINING EFFECTS WITH HARMONICS 14–18 IN THE MID-HIGH STIMULUS

3.1. Stimuli, Procedure and Participants

The stimuli and procedure were the same as for Experiment I, except that for stimulus MID-HIGH the passband extended from 1,400 to 1,800 Hz and the lower slope was 60 dB/oct rather than 30 dB/oct. The spectra of the stimuli are illustrated in Fig. 4. Seven new participants were tested, denoted group MID-HIGH. Training was performed only for the MID-HIGH stimuli, and transfer of learning to the LOW and HIGH stimuli was assessed.

3.2. Results

The left panel of Fig. 5 shows the learning curves. The group mean results showed a progressive improvement across days. Most of the individual results also showed improvements, but with considerable individual variability. A within-subject ANOVA showed a significant effect of session \( F(11, 66) = 2.52, p < 0.001 \). The right panel of Fig. 5 shows the learning and transfer effects. There was a large learning effect for the MID stimuli, with strong transfer to the LOW stimuli and no transfer to the HIGH stimuli. An analysis of variance was conducted on the ratios shown in Fig. 5, right panel. There was a significant effect of stimulus type; \( F(2, 6) = 9.63, p < 0.002 \). Post hoc tests showed that the training effect was smaller for stimulus HIGH than for stimuli LOW or MID-HIGH (both \( p < 0.05 \)), but the effect did not differ for stimuli LOW and MID-HIGH.
The pattern of the results is the same as for Experiment I, despite the fact that all harmonics for stimulus MID-HIGH would have been unresolved. This supports the interpretation that the transfer of learning between stimuli MID and LOW (in Experiment I) and MID-HIGH and LOW (Experiment II) reflects a common underlying mechanism based on the use of TFS information.

4. EXPERIMENT III: TRAINING EFFECTS FOR RANDOM-PHASE TONES

4.1. Rationale

Experiments I and II were conducted using stimuli whose components were added in cosine starting phase (90°), which leads to a waveform on the basilar membrane with a high peak factor when the components are not resolved.

Experiment III was conducted to assess the importance of the peak factor. Components were added with random starting phase. This leads to a waveform on the basilar membrane with a lower peak factor than for cosine phase when the components are not resolved. Effects of component phase on F0 discrimination should only occur when the components on which discrimination is based are at least partly unresolved, so the results were intended partly as a check of whether the components for stimulus MID-HIGH were unresolved.

4.2. Stimuli, Procedure and Participants

The stimuli and procedure were the almost same as for Experiments I and II. For cases LOW, MID-HIGH and HIGH, the passbands extended from 100 to 500, 1,400 to 1,800, and 2,800 to 3,200 Hz, respectively. Twelve new participants were tested. Participants were divided into three groups of four participants, designated LOW, MID-HIGH and HIGH, according to the stimuli used during training. Table 2 lists the harmonics present in the passband during pre/post-testing and training for each group.

4.3. Results

The left panel of Fig. 6 shows the mean learning curves for each group. There was a clear improvement across sessions for groups LOW and MID-HIGH, but not for group HIGH. An analysis of variance with group as a between-subject factor and session as a within-subject factor showed a significant effect of session for groups LOW \( F(11,33) = 2.60, \ p < 0.05 \), and MID \( F(11,33) = 2.73, \ p < 0.01 \), but not for group HIGH \( F(11,33) = 0.61, \ p > 0.05 \).

### Table 2

<table>
<thead>
<tr>
<th>Session</th>
<th>Numbers of the lowest (J) and highest (K) harmonics in the passband (J, K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group LOW</td>
<td>(1, 5)</td>
</tr>
<tr>
<td>Group MID-HIGH</td>
<td>(14, 18)</td>
</tr>
<tr>
<td>Group HIGH</td>
<td>(28, 32)</td>
</tr>
</tbody>
</table>

Fig. 5 The left panel shows F0DLs for the pre-training session (Pre), the training sessions, and the post-training session (Post), for each participant (open symbols) and for the mean (filled circles). The right panel shows learning and transfer effects for each stimulus type. Error bars indicate ±1 standard error.
The mean F0DLs for group HIGH were markedly higher than were obtained for group HIGH in Experiment I, indicating that F0DLs based on temporal envelope cues are affected by the peak factor of the waveform on the basilar membrane, which is consistent with previous work [4,29]. The mean F0DLs for group MID-HIGH were higher than obtained for group MID-HIGH in Experiment II, confirming that the components in stimulus MID-HIGH were at least partially unresolved. The right panel of Fig. 6 shows the training and transfer effects. Group LOW showed a large training effect for the LOW stimuli, with some transfer to the MID-HIGH stimuli and no transfer to the HIGH stimuli. Group MID-HIGH showed a large training effect for the MID-HIGH stimuli, with some transfer to the LOW stimuli and no transfer to the HIGH stimuli. Group HIGH showed no training effect and no transfer to either of the other stimuli. The pattern of the learning and transfer effects is similar to that for Experiments I and II.

An analysis of variance was conducted on the ratios shown in the right panel of Fig. 6, with group membership as a between-subject factor and stimulus type as a within-subject factor. There was a significant effect of group; $F(2, 6) = 14.5, p < 0.001$. Post hoc tests showed that the learning effect was smaller for group HIGH than for groups LOW or MID-HIGH (both $p < 0.002$), but the effect did not differ for groups LOW and MID-HIGH. The effect of stimulus type was significant; $F(2, 24) = 16.01, p < 0.001$. Post hoc tests showed that the learning effect was smaller for stimulus HIGH than for stimuli LOW or MID-HIGH (both $p < 0.05$). The learning effect did not differ significantly for stimuli LOW and MID-HIGH. The interaction of group and stimulus type was significant; $F(4, 24) = 9.71, p < 0.001$. This reflects the fact that the learning effect did not differ across stimulus type for group HIGH, but it did differ across stimulus type for the other two groups.

5. DISCUSSION

The data are partly consistent with those of Grimault et al. [18], in that they found a training effect on F0DLs for complex tones with low resolved harmonics, resembling those for our LOW stimulus. However, Grimault et al. also found a training effect for stimuli with only high unresolved harmonics, resembling our HIGH stimulus, whereas we found no such training effect. The difference may have occurred because our participants were tested using more trials during the pre-training sessions, which would have been sufficient to overcome procedural learning effects. The training effects found by Grimault et al. might have partly reflected procedural learning. Our results also differ from our earlier results obtained for F0 discrimination of a group of harmonics embedded within harmonics whose F0 was fixed [20]. For the earlier results, a training effect was found when the group of components to be discriminated contained only high harmonics, but such an effect was not found here. The training regimes were similar for the earlier study and the present study. The difference across studies therefore probably reflects the difference in the stimuli: discrimination of the F0 of a group of harmonics embedded within harmonics whose F0...
was fixed in the earlier study, versus discrimination of a group of harmonics presented in TEN noise in the present study. In the earlier study, the learning may have involved reduction of the interference effect produced by the background harmonics, an effect known as pitch discrimination interference [30,31].

The present study included stimuli (MID and MID-HIGH) with intermediate harmonic numbers, which were not included in the study of Grimault et al. [18]. In Experiment I, the lowest harmonic within the passband was the 11th, and the lowest component that was above threshold in the TEN was the 7th. In Experiment II, the lowest harmonic within the passband was the 14th, and the lowest harmonic that was above threshold in the TEN was the 11th. It seems likely that only harmonics up to the 8th are resolvable [6,7,10,11,21]. Even if harmonics up to the 10th are resolvable [12], the audible harmonics in the MID-HIGH stimuli were almost certainly only unresolved. Consistent with this, F0 discrimination of the MID-HIGH stimuli was better when the components were added in cosine phase (Experiment II) than when they were added in random phase (Experiment III). The results showed clear training effects for the LOW, MID and MID-HIGH stimuli, and these effects transferred; training with LOW stimuli led to better F0 discrimination of MID stimuli, and training with MID or MID-HIGH stimuli led to better discrimination of LOW stimuli. No training effect occurred for the HIGH stimuli, and training with LOW or MID stimuli did not lead to improved discrimination of HIGH stimuli. F0 discrimination of MID-HIGH and HIGH stimuli was better for stimuli with components added in cosine phase than for stimuli with components added in random phase, consistent with the idea that the components on which discrimination was based were largely unresolved for both MID-HIGH and HIGH stimuli.

We conclude that F0 discrimination is based on similar mechanisms, perhaps based on the use of TFS information, for stimuli with low and intermediate harmonics. A different mechanism, probably based on the discrimination of envelope cues, may be involved for stimuli with only very high harmonics.

ACKNOWLEDGEMENTS

The work of author BM was supported by the Medical Research Council UK (grant number G0701870). We thank two reviewers for helpful comments.

REFERENCES


Hiromitsu Miyazono received M.A. in 1992, and received Dr. Eng. in 2003 at Kumamoto University. He is current professor of Faculty of Administration in the Prefectural University of Kumamoto. His research interest lies in pitch perception, time perception and perceptual learning based on psychoacoustics. He is a member of the Acoustical Society of Japan, Information Processing Society of Japan and the Institute of Electronics, Information and Communication Engineering.

Brian Moore is Professor of Auditory Perception in the University of Cambridge. His research covers both basic and applied aspects of normal and impaired hearing, including hearing aid design and fitting. He is a Fellow of the Royal Society, the Academy of Medical Sciences, and the Acoustical Society of America, and an Honorary Fellow of the Belgian Society of Audiology and of the British Society of Hearing Aid Audiologists. He is President of the Association of Independent Hearing Healthcare Professionals (UK). He is an Associate Editor of the Journal “Hearing Research.” He has written or edited 17 books and over 600 scientific papers and book chapters. His most recent books are “Cochlear Hearing Loss” published by Wiley in 2007 and “An Introduction to the Psychology of Hearing, Sixth Edition,” published by Emerald in 2012. In 2003 he was awarded the Acoustical Society of America Silver Medal in physiological and psychological acoustics. In 2004 he was awarded the first “International Award in Hearing” from the American Academy of Audiology. He has twice been awarded the Littler Prize of the British Society of Audiology. In 2008 he received the Award of Merit from the Association for Research in Otolaryngology and the Hugh Knowles Prize for Distinguished Achievement from Northwestern University.