Acoust. Sci. & Tech. **22**, 3 (2001)

PAPER

The relationship between the fluctuations of harmonics and the subjective quality of flute tone

Akira Nishimura[1], Mitsumi Kato[2] and Yoshinori Ando[1]

[1]*Department of Information Systems, Faculty of Business Administration and Information Science, Tokyo University of Information Sciences,*
1200–2, Yato-cho, Wakaba-ku, Chiba, 265–8501 Japan
[2]*Faculty of Music, Kurashiki Sakuyo University,*
3515, Tamashimanagao, Kurashiki, 710–0292 Japan

(*Received 10 May 2000, Accepted for publication 28 February 2001*)

Abstract: We studied the relationship between amplitude and frequency fluctuations of harmonics and the perceived quality of flute tones with vibrato. To investigate the effects of minute and irregular fluctuations on timbre, a real flute tone and synthesized flute tones whose relative amplitude levels of harmonics and extent of vibrato were equal to those of the real tone, were used for the subjective experiments. Listener's preference for flute tones was found to be affected by the degree of intensification or attenuation of the frequency and amplitude fluctuations above 13 Hz. Also, we investigated the physical properties of the fluctuations that affect perceived quality of flute tones, by synthesizing fluctuation waves of harmonics. The results of evaluation by test subjects show that there was no perceived difference in quality between the original tone and synthesized tones with fluctuations that were synthesized by randomization of the phase spectra of the original fluctuations. In contrast, synthesized tones with fluctuations that were synthesized from filtered noise were perceived to be significantly inferior to the original tone. These results suggest that spectral variation of fluctuation waves which is at higher frequency and lower amplitude than spectral variation of vibrato influences perceived quality.

Keywords: Flute, Harmonics, Fluctuation, Analysis by synthesis, Subjective quality

PACS number: 43.66.Jh, 43.75.Qr

1. INTRODUCTION

Several studies have been conducted on the relationship between the amplitude envelope or spectral envelope of the sounds of musical instrument and their timbre. Relatively slow and regular fluctuation of harmonics (that is, vibrato) has been shown to be one of the factors which affect the perceived naturalness of timbre of musical instruments. However, few studies have been conducted on the relationship between minute fluctuations of amplitude and frequency (non-vibrato) and the perceived beauty or preference of timbre. A series of studies reported by Ando and his colleagues represent the only known attempts to clarify these relationships. Such minute fluctuations of harmonics have been found to have positive effects on the perceived quality of air-reed instruments [1–3], but they have been found to have either no effect or negative effects on the perceived quality of other wind instruments, such as reed wind instruments [4] and brass wind instruments [5].

Ando and Shima investigated the relationship between the physical properties of sustained flute tones, without vibrato, and their perceived qualities [2]. The flute tones used in their experiment were produced by players whose skill levels ranged from beginner to professional. The tones produced by these players had different amplitude spectra and different degrees of fluctuation in frequency and amplitude. These two factors, the amplitude spectrum and the amount and nature of fluctuations, may both affect the timbre of flute tones simultaneously. For this reason, they conducted an experiment with synthesized flute tones, to investigate the effect of amplitude fluctuations on the timbre of flute tones. They varied the amplitude modulation ratio of the harmonics, while keeping the amplitude spectrum constant. In their experiment, the listener's judgment of tone quality varied depending on the extent of their experience in flute playing. The subjects who had no experience in flute playing evaluated the flute tone with no amplitude modulation as superior in quality. In contrast, one of the present authors who had experience playing the flute preferred the flute tone with moderate fluctuations.

Ando *et al*. applied almost the same method to the other air-reed instruments, the treble recorder [6] and the syakuhachi [3], and they reported that professional players preferred tones which contained moderate fluctuations in their harmonics.

In addition to these studies, Ando and Sakagami investigated the relationship between fluctuation characteristics and the timbre of sustained treble recorder tones [7]. They conducted experiments to evaluate similarities in timbre between real recorder tones and synthesized recorder tones, keeping the amplitude spectrum constant. Their results showed that a tone synthesized from frequency and amplitude fluctuation waves extracted from each harmonic of a real tone had a timbre close to that of the real tone. In contrast, the synthesized tone whose harmonic components were composed of pure tones, whose other components were composed of random noise, and whose long-term spectrum was equated to the original tone had a different timbre than the original tone and the synthesized tone generated from the original fluctuation waves.

The general conclusion that can be drawn from these findings is that not only the amplitude spectrum (which is regarded as the most important factor for timbre of sustained musical tones) but also the frequency and amplitude fluctuations of harmonics play an important role in the quality of air-reed instrument tones. However, previous studies have only dealt with musical tones without vibrato. Therefore, it is not obvious whether the findings obtained from those studies will hold true for flute tones played with vibrato. In addition, the results of informal experiments by Ando indicate that more minute properties of fluctuation waves probably affect the quality of air-reed instrument tones.

Here, we are concerned with flute tones with vibrato, and we focus on the amplitude and frequency fluctuations of their harmonics. The target frequency range of these fluctuations is above the vibrato frequency, because we want to investigate minute and nonperiodic fluctuations that a flute tone intrinsically contains. In this paper, we report on our investigation into the relationship between these amplitude and frequency fluctuations of harmonics and the perceived quality of flute tones.

2. FLUCTUATION ANALYSIS OF HARMONICS OF FLUTE TONES

Analysis of air-reed instrument tones in the previous studies involved sustained tones without vibrato. However, in real musical performance, sustained tones often contain vibrato. In the experiment described in this section, we analyzed sustained flute tones with vibrato to investigate physical properties of frequency and amplitude fluctuations of their harmonics.

2.1. Sound Material

Flute sounds were recorded on a DAT recorder at a sampling frequency of 48 kHz, with a microphone (B&K 4190) positioned 50 cm from the mouth hole. The recorded melody was the cadence composed by Donjon from Mozart's "Concert No. 2 for Flute and Orchestra in D K.314," played by Prof. Iwasaki of Kurashiki Sakuyo University, Faculty of Music. The recorded sounds were down-sampled to 22.05 kHz before analysis. The sound which was analyzed was the first note of the melody, and its duration was 2.1 seconds. Its fundamental frequency was approximately 900 Hz.

2.2. Frequency Analysis of Fluctuation Wave

Harmonic components were separated by a bandpass filter bank whose bandwidth was the fundamental frequency and whose center frequencies were integer multiples of the fundamental. Frequency and amplitude fluctuation waves of each harmonic were then extracted using analytic signals. The analytic signal is a complex signal in which the real component is the original signal and the imaginary component is its Hilbert transform. The frequency fluctuation wave is obtained by the derivation of the phase of the analytic signal. The amplitude fluctuation wave is derived from the absolute value of the analytic signal.

Figure 1 shows the frequency fluctuation waveforms of the first to sixth harmonics, relative to the integer multiples of the fundamental frequency. In the figure, the harmonics are sequentially shifted apart by −10%, for clarity. Figure 2 shows amplitude fluctuation waveforms of the first to sixth harmonics, relative to the maximum

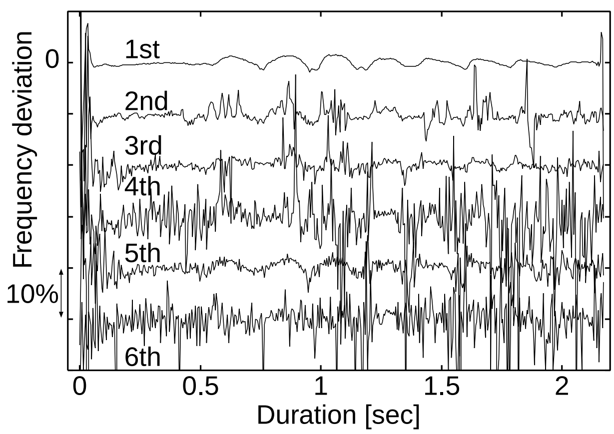

Fig. 1 Frequency fluctuation waveforms of the first to sixth harmonics relative to the integer multiples of the fundamental frequency. The harmonics are sequentially shifted apart by −10%, for clarity.

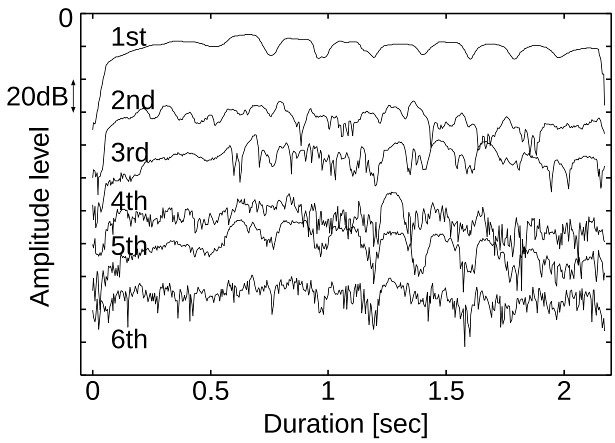

Fig. 2 Amplitude fluctuation waveforms of the first to sixth harmonics relative to the maximum amplitude level of the sound file format. The harmonics are sequentially shifted apart by −20 dB, for clarity.

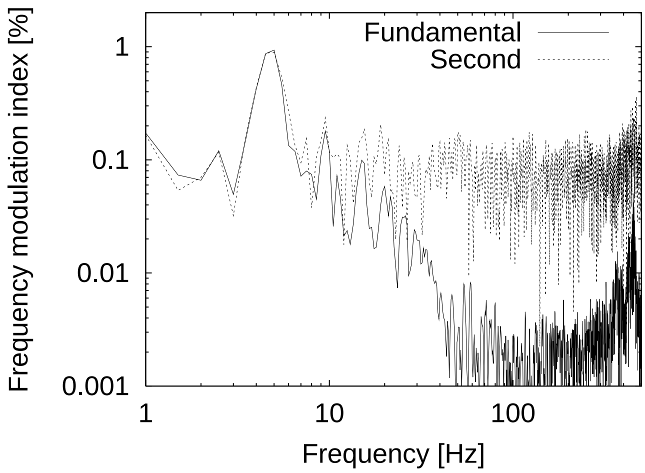

Fig. 3 Power spectra of frequency fluctuation extracted from the fundamental and the second harmonic.

amplitude level of the sound file format (Microsoft RIFF audio file). Here, the harmonics are sequentially shifted apart by −20 dB, for clarity.

Long-term spectrum analysis [44,100-points discrete Fourier transform (DFT) after 44,100-points Hanning-windowing] was applied to frequency and amplitude fluctuation waves of each harmonic.

Figure 3 shows the spectra of frequency fluctuation waves extracted from the fundamental and the second harmonic, as an example. All harmonics show vibrato at approximately 5 Hz with modulation index of approximately 1%. In the fundamental, strong harmonic components of vibrato are seen in the frequency range below 30 Hz. However, in the frequency range 30 Hz to 400 Hz, only slight fluctuations of approximately 0.01% in modulation index are observed. In the second harmonic, the strong component is only seen at vibrato of 5 Hz, and at other frequencies the modulation index remains constant at approximately 0.1%. The spectra of the frequency fluctuations extracted from the harmonics above the second harmonic (not shown) show the same spectral envelope as the second harmonic. The results of the spectral analysis of the other tones extracted from the recorded melody show almost the same tendency as the first tone. The amount of frequency fluctuation, other than vibrato, is almost the same as the value reported by Kunii and Ando [1] (less than 0.3%) for flute tone without vibrato.

Figure 4 shows power spectra of the amplitude fluctuation waves extracted from the fundamental to the sixth harmonic. Each harmonic shows a constant negative slant in its spectrum. These slants are almost the same as for the results obtained for the other tones from the recorded melody. The slant in the spectrum of the fundamental is −10 to −15 dB/octave. The degree of slant gradually decreases with increasing harmonic number, and it becomes −3 dB/octave above the 6th harmonic. Ando and Shima reported octave-band spectra of the amplitude fluctuation waves extracted from flute tones without vibrato. They reported that the slant in the spectra in the central frequency range of 6.4 Hz to 101.6 Hz gradually decreased with increasing harmonic number; it was −7 dB/octave for the fundamental and −3 dB/octave for the 6th harmonic, though individual differences among the samples of each note were observed. We calculated octave-band levels identical to theirs, using amplitude fluctuation waves extracted by a method in which the peak value of every cycle of the output waveform of the harmonic bandpass filterbank was determined by quadratic approximation. Our results show a slant in the spectrum of the fundamental which is steeper (−15 dB/oct.) than that ob-

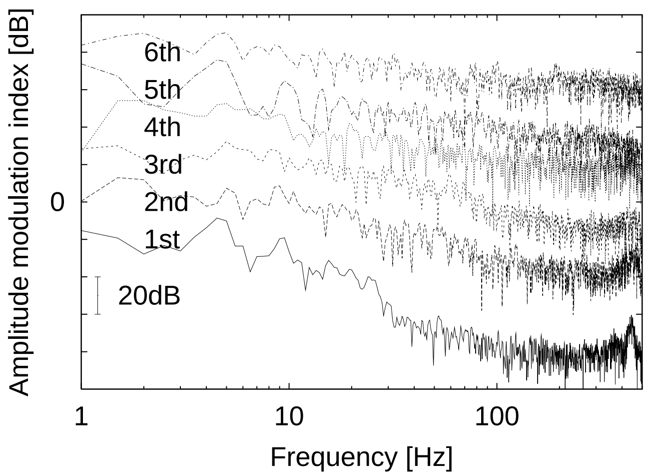

Fig. 4 Power spectra of amplitude fluctuation extracted from the fundamental to sixth harmonic. The harmonics are sequentially shifted apart by 20 dB, for clarity.

tained by Ando and Shima, although the aspects of the spectral envelope of the other harmonics that we observed are in line with their results. The difference between these two results for the fundamental may be due to the effects of the vibrato component in the region of low fluctuation frequency.

3. EXPERIMENT 1

A preliminary experiment was conducted to evaluate the perceived quality of synthesized flute tones with intensified or attenuated frequency and amplitude fluctuations of harmonics, with the vibrato fluctuation and the amplitude spectrum held constant. The results showed that the perceptions of timbre differed according to variations in the power of fluctuations. Therefore, we attempted to determine which frequency region of the fluctuations affects the perceived quality of flute tones, by attenuating or intensifying amplitude and frequency fluctuations of each harmonic.

3.1. Stimuli

The flute tone used in this experiment was the same as in Section 2: the first note of the melody. Its duration was approximately 2.1 seconds. Table 1 shows the parameters for synthesis. The meanings of the labels of the different stimuli are as follows. The stimulus orig is the original tone. The stimulus s_peak was synthesized from a time sequence of amplitude and frequency data which were calculated by 3-point interpolation [8] around spectral peaks obtained from 128-point fast Fourier transform (FFT), shifting 64 points each. The letters used in the label of the other stimuli have the following meanings: 'f', frequency fluctuation; 'a', amplitude fluctuation; 'w', wide band; 'h', high boost condition. The numbers represent cutoff frequency of the fluctuation filtering. Thus, the stimulus fwa13 was synthesized using the frequency fluctuation waves extracted from the harmonics of the original tone and the amplitude fluctuation waves which were lowpass-filtered at 13 Hz. Characteristics of the fluctuation filters are shown in Fig. 5. All filters left the vibrato frequency region unchanged. All stimuli except orig were synthesized up to the 10th harmonic.

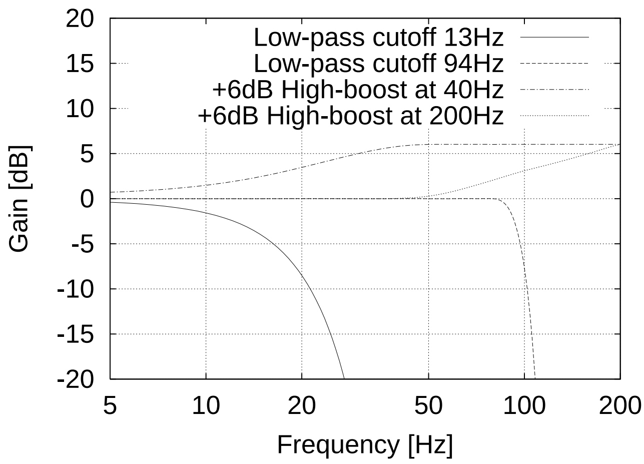

Fig. 5 Characteristics of filters.

Average power spectra of the nine stimuli, calculated by half-overlapped 4,096-point FFT with Hanning-windowing, are shown in Fig. 6. The apparent noise floor, at a level of -60 dB relative to the fundamental, is not background noise. It represents inherent noise that every flute tone has. No matter how well a flautist may play, or how silent the room is, these noise components will appear in spectra. A subharmonic series of order two which was observed in the original tone and the synthesized tones reflects the first resonance frequency of a flute pipe.

Table 1 Stimuli and their parameter for synthesis in experiment 1.

label	method of synthesis	
	filter of frequency fluctuation	filter of amplitude fluctuation
f13a13	low-pass 13 Hz	low-pass 13 Hz
f94a94	low-pass 94 Hz	low-pass 94 Hz
s_peak	128 points FFT, 64 points overlapped (Sinusoidal modeling synthesis)	
f13aw	low-pass 13 Hz	—
fwaw	—	—
fwa13	—	low-pass 13 Hz
orig	original	
fwah200	—	+6 dB high-boost at 200 Hz
fwah40	—	+6 dB high-boost at 40 Hz

3.2. Experimental Procedure

The experiment consisted of three listening sessions per subject. The nine stimuli shown in Table 1 were presented as pairs of stimuli, and all possible combinations including reverse order were presented to each test subject once per session, in random order. The subjects answered the preference of the latter tone compared to the former tone, in five steps, immediately after the pair was played. The total number of the pairs presented in a session is 72. Details of the nine subjects are as follows: Subjects 1 and 5 were two of the present authors; Subjects 2, 3 and 4 were amateur flute players; Subjects 6, 7, 8 and 9 were undergraduate students of the faculty of music, majoring in flute. Stimuli were presented diotically via headphones (Audio-Technica ATH-A5X). Sound pressure levels [measured by artificial ears (B&K 4153 and B&K 4134)] were in the range from 63 to 69 dB. The results of each subject's first session were not used for the anal-

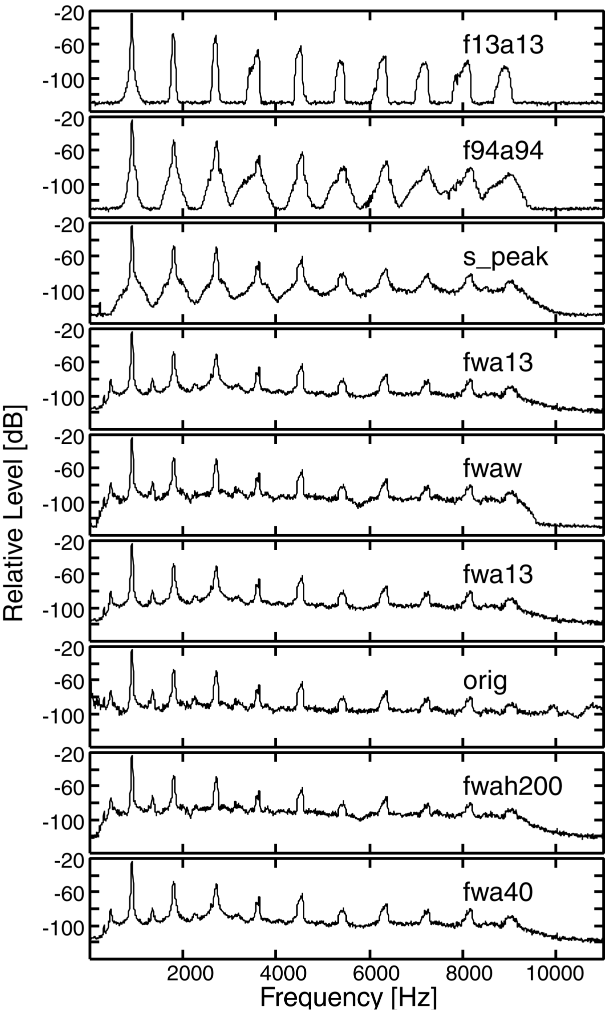

Fig. 6 Average power spectra of stimuli.

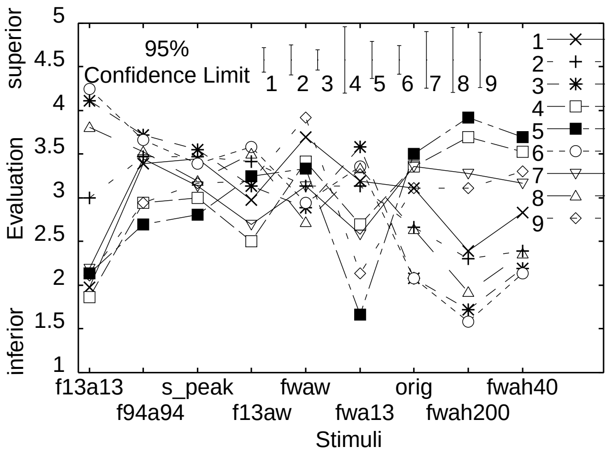

Fig. 7 Estimated preferences for flute tones with 95% confidence limit.

ysis; this was meant to compensate for results which may have been due to subjects' lack of familiarity with the experimental procedure.

3.3. Results

Scheffé's method for analysis of variance (ANOVA), modified by Ura, was applied to the results of the experiments. Estimated values of preference, with their 95% confidence limits, are shown in Fig. 7.

A cluster analysis was performed on the subjects. The subjects were classified into two groups, one consisting of subjects 4, 5, 7 and 9, and the other consisting of subjects 2, 3, 6 and 8. All Spearman's correlation coefficients obtained within each group of subjects were significant ($p < 0.05$). The former group tended to prefer flute tones with more fluctuation in the harmonics, and the latter group tended to prefer tones with less fluctuation.

Ando and Shima calculated average levels of components of flute tones without vibrato in the frequency range 2 kHz to 5 kHz excluding 100 Hz bandwidth centered around the harmonic frequencies. They assumed that these noise components were not due to inherent fluctuation of the harmonics, but rather to inherent noise. They found that the perceived quality of a real flute tone deteriorated with increases in this noise level. We also calculated the average noise levels of the stimuli, and investigated the relationship between these noise levels and listener's preference. As we show in Fig. 3, the dominant component of the vibrato fluctuation had a frequency of 5 Hz and modulation index of 1%; these values are almost constant among all harmonics. We observed no periodic amplitude or frequency fluctuations, other than vibrato, in any of the harmonics. Therefore, in this case, it is appropriate to consider fluctuation components other than vibrato fluctuation to be random fluctuations, and these fluctuations generate noise components. However, the side band components generated by the vibrato fluctuation may effect outside the 100 Hz bandwidth centered around the harmonic frequencies. Therefore, we calculated the power levels of the side band components which are generated by the frequency modulation of the vibrato which has a frequency of 5 Hz and a modulation index of 1%. The power level of the 5,300 Hz component, which is 100 Hz lower than the carrier frequency of 5,400 Hz (the 6th harmonic), is -73 dB relative to the carrier frequency's level. This result means that it is appropriate to neglect the effects of the side band components generated by the vibrato fluctuation on the spectral components with frequencies more than 100 Hz distant from the harmonic components. Therefore, we regarded the average level of the components in the frequency range from 1,900 to 5,300 Hz, excluding the 200 Hz bandwidth centered around the harmonic frequencies, as the noise level, and calculated this level for each stimulus from the results of the spectral analysis (Fig. 6). Figure 8 shows the

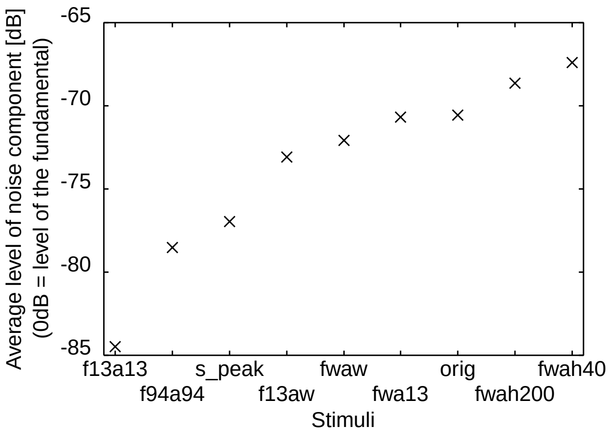

Fig. 8 Average noise level.

average noise levels for all stimuli.

Correlation coefficients between these noise levels and estimated values of preference were calculated for each subject. Negative correlation was significant ($p < 0.05$) for Subjects 3, 6 and 8. According to these results, some of our subjects had a tendency to prefer flute tones with low noise levels, like the amateur subjects who participated in the experiment conducted by Ando and Shima [2]. Figure 7 reveals that all subjects showed significant differences ($p < 0.05$) in preference between f94a94 (in which fluctuations above 94 Hz were eliminated) and fwaw (which consisted of the original fluctuations). This result demonstrates that fluctuations above 94 Hz affect the perceived quality of flute tone, although the extent of preference for the tones varied by subject. Moreover, eight of our nine subjects showed significant difference in preference between f94a94 and f13a13, and all subjects showed significant difference in preference between fwaw and f13a13 ($p < 0.05$). Therefore, fluctuation components at frequencies from 13 to 94 Hz can also affect the perceived quality of flute tones.

3.4. Discussion

3.4.1. Relationship between subject's playing experience and preferred amount of fluctuation

In a series of studies by Ando and his colleagues [2, 3, 6] which dealt with air-reed instruments, professional players were found to prefer those synthesized tones which had moderate degrees of fluctuation, whereas amateur players preferred synthesized tones with lesser degrees of fluctuation.

In the present experiment, four of the nine subjects were undergraduate music students majoring in flute, and they were considered to be more experienced flute players than the other subjects. Two of the four subjects who preferred tones with lesser degrees of fluctuation, such as f13a13, were undergraduates majoring in flute. The other two undergraduates majoring in flute preferred tones with moderate fluctuation, such as fwaw or orig. These results suggest that, among our subjects, the preferred amount of fluctuation was not affected by the level of flute playing experience.

In the present experiment, the synthesized flute tone with a low degree of fluctuation was f13a13. This tone differs from the synthesized flute tone with no fluctuations, which was the one most preferred in Ando and Shima's experiment [2], in that it contains low-frequency fluctuations (below 13 Hz) and vibrato. This difference may be the source of the disagreement seen here between the preferences reported in the present study and the previous study.

3.4.2. The nature of the Fluctuations and Spectra of Synthesized Tones

There were two pairs of stimuli that had nearly equal average noise levels (Fig. 8) but produced noticeably different results for preference (Fig. 7). The first pair was fwa13 and orig, which had average noise levels that differed by only 0.1 dB. The other pair was f13aw and fwaw, which had average noise levels that differed by 1.0 dB. Six subjects showed significant differences in preference for both pairs ($p < 0.05$).

The noise components with frequencies outside of the 200 Hz bandwidth centered around the harmonic frequencies were considered to be the side band components generated by the amplitude fluctuations (above 100 Hz) and the frequency fluctuations (excluding the vibrato components) of the harmonics. Although the energy levels of these noise components were nearly the same for all the stimuli in this study, the nature of the fluctuations which generated them differed, as follows: only amplitude fluctuation (f13aw), both amplitude and frequency fluctuations (fwaw, orig), only frequency fluctuation (fwa13). These differences may affect the quality of timbre, regardless of whether the side band components which are generated by the fluctuations have equal levels of power.

4. EXPERIMENT 2

In Experiment 1, we demonstrated that fluctuation components with frequencies above 13 Hz significantly influence the quality of flute tone. Also, it was found that there are "physical correlates" of the flute tone in the fluctuation frequency region. The term "physical correlates" means the physical properties that affect listener's evaluation. The most likely physical correlate, which was dealt with in previous studies and investigated in our analysis of fluctuations (see Section 2), is the spectral envelope of the fluctuation wave. In this section, we report on the results of our investigation into the physical properties of the fluctuations that affect listener's evaluation of flute tone, using a method involving synthesis of fluctuation waves of harmonics.

4.1. Characteristics of Fluctuation Waves

Ando and Shima [2] synthesized amplitude fluctuation waves below the 7th harmonic, using octave-band noise that was equal to the octave-band power levels of the amplitude fluctuation waves of a flute tone which was judged by listeners to be of excellent quality. However, the resultant synthesized flute tone and the original flute tone were not compared in a listener's preference test, so we do not know what effects their method of synthesizing amplitude fluctuation waves from noise had on perceived tone quality. If there was no difference in perceived quality, the amplitude fluctuation of each harmonic may have had properties similar to those of filtered noise with a spectral envelope similar to that of the original tone.

As we discussed in Section 2, for all the recorded tones, the long-term spectra of the frequency and amplitude fluctuation waves of the harmonics (Fig. 3 and Fig. 4, respectively) are nearly the same as those of the first tone of the melody. However, as we reviewed in Section 1, synthesized recorder tones with harmonic components composed of pure tones, other components composed of noise, and long-term spectra equated to the original tone had different timbre than the original tone [7]. One of the reasons for this timbre difference may be that temporal variation of fluctuation spectra was not taken into account.

Thus, we calculated spectrograms of the fluctuation waves of the harmonics, to observe their temporal and spectral variations. The fluctuation spectrogram was calculated by FFT with 1,024-point Hanning-windowing, shifting every 256 points.

Figure 9 shows temporal variation of the fluctuation components at specific fluctuation frequencies. These components were extracted from the amplitude fluctuation spectrogram of the third harmonic of the first note of the melody. The vertical axis represents relative amplitude level (relative to the maximum amplitude level of the sound file format). The DC (direct current) component is the lowest-frequency (below 11 Hz) component of the fluctuation spectra; i.e., temporal variation of the vibrato component. Fluctuation components at 215 Hz and 323 Hz are roughly synchronized to the power of the DC component. The other fluctuation frequency components show the same tendency as these fluctuation components.

These results indicate that the power levels of the fluctuation components above the vibrato frequency are roughly synchronized to that of the vibrato. In other words, the fluctuation components above the vibrato frequency are modulated by the vibrato component. However, this synchronization was not observed in some other tones in the melody. As we showed in Fig. 3 and Fig. 4, the fluctuation waves did not contain periodic components, with the exception of the vibrato component. These findings mean that the amplitude and frequency fluctuation waves with frequencies above the vibrato frequency are amplitude-modulated random signals under certain conditions.

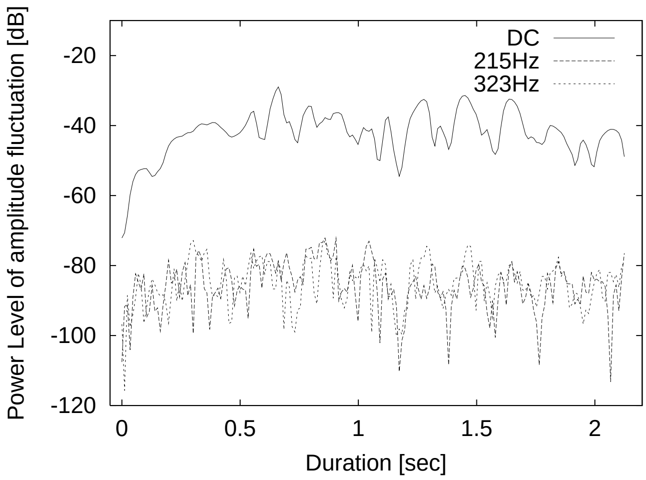

Fig. 9 Temporal variations of the amplitude fluctuation power at DC (direct current), 215 Hz and 323 Hz. Fluctuation components at 215 Hz and 323 Hz are roughly synchronized to the power of DC components.

Thus, we developed a method for synthesizing fluctuation waves which include temporal and spectral variations. We then conducted an experiment designed to investigate the relationship between the spectral properties of fluctuation waves and the perceived quality of flute tones.

4.2. Outline of the Method for Analysis Using Synthesized Fluctuation Waves

The algorithm for synthesis of amplitude and frequency fluctuation waves of harmonics is illustrated in Fig. 10. To analyze and synthesize sustained parts of tones, input sound was separated into individual notes on the basis of the fundamental frequency transition. Before analysis, we selected which notes would be synthesized. The notes not synthesized were appended to the synthesized notes. The initial attack and final decay parts of the notes were not synthesized, and those of original waves were appended to the synthesized sustained tone. Harmonics were separated up to the 10th harmonic using a filter bank, and the amplitude and frequency fluctuation waves of each harmonic were extracted using the analytic signal. Components above the 10th harmonic were not

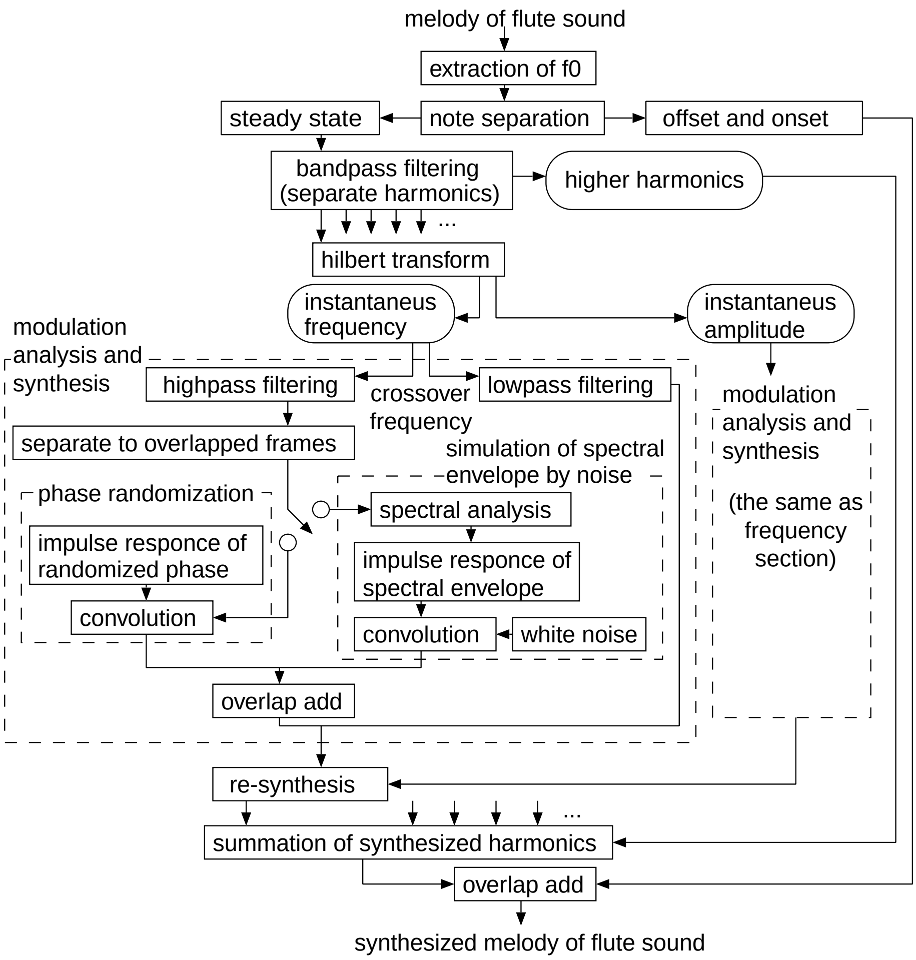

Fig. 10 Algorithm of synthesizing amplitude and frequency fluctuations.

modified and they were added to the synthesized harmonics.

The amplitude and frequency fluctuation waves of each harmonic were separated by low-pass and high-pass filters, the crossover frequencies of which were specified as parameters for synthesis. Lower-frequency fluctuation waves were not modified and they were added to the higher-frequency synthesized fluctuation waves, so that the vibrato fluctuation would remain unchanged. Analysis using synthesized fluctuation waves was carried out for higher-frequency fluctuation waves. All filters used in the analysis and synthesis procedures were linear-phased finite impulse response (FIR) filters, and time lags of filtering were all compensated.

4.3. Analysis by Synthesis of Fluctuation Waves

The fluctuation waves above the crossover frequency were separated into frames of 1,024 points which overlapped by 512 points. After Hanning-windowing for each frame, the fluctuation wave of each frame was synthesized by one of the two methods described in Sections 4.3.1 and 4.3.2. Synthesis was applied to either the amplitude fluctuation wave or the frequency fluctuation wave. For the fluctuations which were not synthesized, extracted fluctuation waves were used to synthesize harmonics.

To verify the accuracy of the method of synthesis described above, synthesis was conducted using the original fluctuation waves. The resultant waveform of the synthesized tones was identical to that of the original tones. Therefore, we determined that the methods we used for fluctuation analysis were sufficiently accurate. In addition, since these methods of synthesis only involved modification of the fluctuation components higher than the crossover frequency, average power level and average frequency of the harmonics and vibrato fluctuations were equal to those of the original flute tone.

4.3.1. Synthesis using filtered Gaussian noise

Frequency analysis was performed with a FFT, to obtain the power spectrum of the waveform of each frame. A 1,024th-order FIR digital filter with a frequency response specified by this power spectrum was designed using the Hanning window method. The impulse response of the filter and the Hanning-windowed Gaussian noise signal were convolved to obtain a noise signal which simulated the spectral envelope of each frame. A Gaussian noise signal was independently generated for every frame.

4.3.2. Synthesis by randomization of Phase spectra

An impulse response with a phase spectrum randomized in the range $-\pi$ to π was calculated using an inverse FFT. This impulse response and each frame's waveform were convolved to obtain a phase-randomized signal; a random phase spectrum was independently calculated for every frame.

4.4. Analysis of Synthesized Fluctuation Waves

To compare both spectral and temporal similarities of the original and synthesized fluctuation waves, two kinds of correlation coefficients were calculated from the spectrogram of the original fluctuation wave and that of the synthesized fluctuation waves with crossover frequency of 100 Hz. The calculation of spectrograms was conducted according to the method described in Section 4.1. The amplitude fluctuation wave extracted from the third harmonic of the first note of the melody was analyzed as an example.

First, investigation was conducted to determine whether synchronization between amplitude variation of the vibrato and that of the fluctuation components (as shown in Fig. 9) was the same for the original tone and the synthesized tones. Specifically, we determined correlation coefficients between temporal variation of the vibrato component and that of the other fluctuation components; i.e., temporal variation observed in the DC component of the spectrogram and that observed in the other discrete fluctuation frequency components of the spectrogram. These correlation coefficients were calculated from the original and synthesized amplitude fluctuation waves (Fig. 11). The fact that the two synthesized fluctuation waves had positive correlation coefficients which were nearly equal to that of the original fluctuation wave indicates that the two methods of synthesizing fluctuation waves accurately simulated the characteristics of the

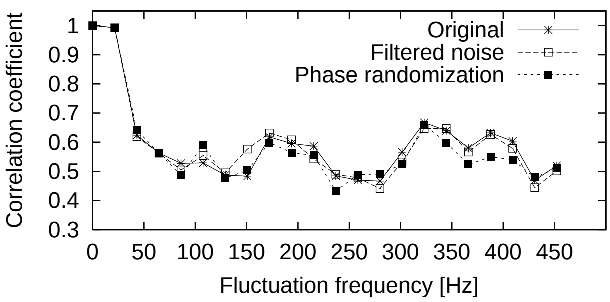

Fig. 11 Correlation coefficients between DC component and the higher frequency components of the amplitude fluctuation spectra of the third harmonic.

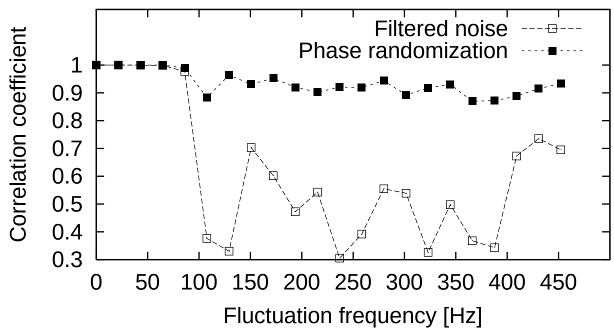

Fig. 12 Correlation coefficients between spectral components of the amplitude fluctuation spectra of the third harmonic observed from the original and that observed from the synthesized.

original amplitude fluctuation wave; i.e., the power of the fluctuation components roughly synchronized to that of the vibrato.

Next, we investigated how accurately minute spectral and temporal variations of the original fluctuation waves were simulated in the synthesized fluctuation waves. Specifically, we calculated correlation coefficients between the temporal variation of the original amplitude fluctuation spectra and that of the synthesized amplitude fluctuation spectra, for each discrete frequency. Figure 12 shows the correlation coefficients obtained from the two methods of synthesizing fluctuation waves. Below approximately 100 Hz, correlation coefficients of both synthesis methods are nearly equal to unity, because original fluctuations below 100 Hz were used for synthesis. Above 100 Hz (crossover frequency), correlation coefficients of the phase randomization method are greater than 0.87, indicating that the minute temporal variations of the spectra were accurately simulated. In contrast, correlation coefficients of the filtered noise method are in the range 0.33 to 0.74, indicating that the filtered noise method resulted in inaccurate simulation of the small temporal variations of the original spectra.

4.5. Experimental Procedure

With subjects who were experienced flute players, we felt it would be better to use a comparatively long melody as the stimulus for evaluation. Therefore, we chose two parts of the recorded flute sounds, the first and second measures and the 15th and 16th measures, as test stimuli. These melodies are shown in Fig. 13. Triangles correspond to the target notes for synthesis, and original tones were used for the other notes.

Three experimental parameters are shown in Table 2: method of fluctuation synthesis (of the two described in Section 4.3), crossover frequency, and the kind of fluctuation which was synthesized. The meaning of the labels

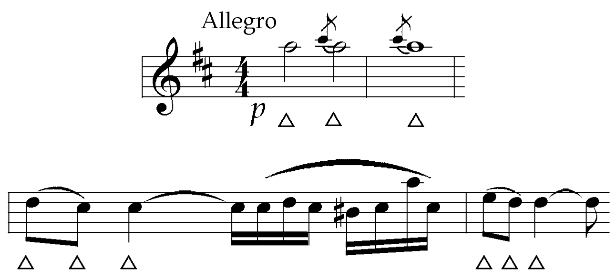

Fig. 13 Top: The first and second measures; Bottom: The 15th and 16th measures. Triangles correspond to the target notes for synthesis.

of the test sounds is explained below: the letters 'a' and 'f' have the same meaning as they do in Experiment 1. The letter 'n' means that the filtered Gaussian noise synthesis method was used in the synthesis of the tones, and 'r' means that the phase randomization method was used in the synthesis of the tones. The numbers '20' and '100' indicate the crossover frequencies.

The experiment consisted of three listening sessions per subject. The nine stimuli shown in Table 2 were presented as pairs of stimuli, and all possible combinations including reverse order were presented to each subject once per session, in random order. The subjects answered the preference of the latter tone compared to the former tone, in five steps, immediately after the pair was played. Details of the 8 subjects were as follows: Subjects AN and MK were two of the present authors; Subject MO was an amateur flute player who was a university undergraduate studying music; Subjects KN, NA, SS and KG were undergraduates majoring in flute; Subject NI was the player who produced the flute sounds used in this experiment. Stimuli were presented diotically via headphones (Audio-Technica ATH-A5X). Sound pressure levels were controlled freely by each subject, within the range of 56 to 74 dBSPL [measured with artificial ears (B&K 4153 and B&K 4134)].

Table 2 Stimuli and their parameters for synthesis in experiment 2.

label	fluctuation	crossover (Hz)	method
orig	(original)	—	—
a20r	amplitude	20	phase randomization
a20n	amplitude	20	filtered noise
f20r	frequency	20	phase randomization
f20n	frequency	20	filtered noise
a100r	amplitude	100	phase randomization
a100n	amplitude	100	filtered noise
f100r	frequency	100	phase randomization
f100n	frequency	100	filtered noise

4.6. Results

Scheffé's method of ANOVA, modified by Ura, was applied to the results of the experiments. Significant interactions between session and stimulus were found for a minority of the subjects. This significant interaction was found in the evaluation of the first and second measures obtained from MO. The same significant interactions were found in those of the 15th and 16th measure obtained from KN and MK. For subject MK, the evaluations of the second session were found to be significantly different from those of the other two sessions. It may be due to a lack of stability of his internal reference for evaluation, because he had no experience on flute playing. For subject MO and KN, the evaluations of the first session were found to be significantly different from those of the other two sessions. It may be due to a lack of familiarity with the stimulus sounds. For these subjects, the session in which the inconsistent evaluation occurred was excluded, and ANOVA was applied again to the data from the other two sessions. In all such cases, there was no significant interaction between session and stimulus. Therefore, with this modification, we regarded the data from these subjects as reproductive data, and we used them in our analysis.

Estimated values of preference obtained from the experiments using the first and second measures are shown in Fig. 14 for all subjects, except NA and KN. NA and KN showed no significant difference in preference for any of the stimuli, so their data were excluded from Fig. 14 and the subsequent analysis. The results obtained from the experiments using the 15th and 16th measures agreed with those obtained for the first and second measures, so that data is not included in the figures.

Figure 14 illustrates that the tendencies of preference were very similar among all subjects, although their

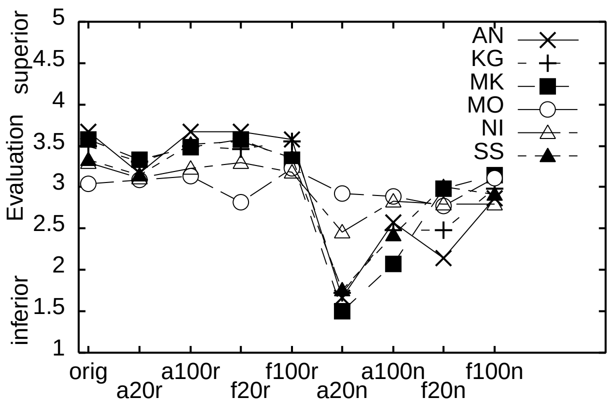

Fig. 14 Estimated values of preference. Evaluated sounds were performance of the first and second measures.

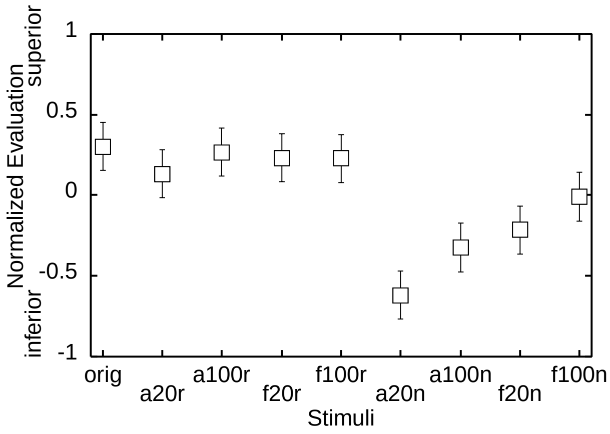

Fig. 15 Average value of normalized preference with 95% confidence limits. Evaluated sounds were performance of the first and second measures.

ranges of estimated values were different. For example, the range of estimated values was less than unity for subjects NI and MO, whereas that of the other subjects was approximately 2. These differences are due to the ranges of evaluation that the subjects used. The subjects NI and MO answered using small ranges of evaluation, even though they were clearly able to perceive timbre differences.

To equate the ranges of evaluation, the data obtained from each subject were normalized using the difference between the maximum and minimum values of their estimations as a standard. ANOVA was again applied to the data after normalization.

Figure 15 shows estimated preferences averaged over all subjects. Error bars correspond to 95% confidence limits. An interaction between stimulus and subject was not significant [$F(128, 1043) = 0.93$ ($p < 0.697$)]. This finding means that no subject evaluated any of the stimuli differently than did the other subjects.

Figure 15 also shows that there was no significant difference in perceived quality between the synthesized flute tones which were generated by the phase randomization method (f20r, f100r, a20r, a100r) and the original flute tone (orig). This finding means that the phase of the fluctuation waves of the harmonics did not affect the perceived quality of these flute tones. In contrast, the synthesized flute tones which were generated by the filtered noise method (a20n, a100n, f20n, f100n) were evaluated as significantly inferior in quality to the original tone ($p < 0.05$). Both methods of fluctuation synthesis maintain temporal variations of the spectra which synchronize to the amplitude variation of the vibrato, as shown in Fig. 11. However, temporal variation of the power of each fluctuation frequency component which was generated by the phase randomization method is closer to that of the original fluctuation component (see Fig. 12). These results indicate that spectral variation of the fluctuation waves which is higher in frequency and lower in amplitude than the spectral variation which accompanies vibrato may have an influence on perceived quality.

In the case of the filtered noise synthesis method, those tones with synthesized frequency fluctuation waves (f20n, f100n) were evaluated as being of not as poor quality as the tones with synthesized amplitude fluctuation waves (a20n, a100n). This result means that frequency fluctuation waves are less likely to lose physical correlates to timbre than amplitude fluctuation waves. Also, the timbre of the flute tones synthesized using high crossover frequency (a100n, f100n) was evaluated as significantly superior to that of tones synthesized using low crossover frequency (a20n, f20n). Since the frequency range of fluctuation synthesis is wide when the crossover frequency is low, the crossover frequency may affect the quality of the synthesized tone. This result indicates that not only the existence of higher fluctuation components but also the characteristics of fluctuation components in the frequency range of 20 Hz to 100 Hz and of above 100 Hz affect the timbre of flute tones.

4.7. Discussion

In Experiment 2, we investigated the effects of the modification of fluctuation waves on timbre of flute tones. We assume that those physical properties whose presence or absence affects perceived quality are physical correlates of flute tone. However, it is possible that other physical correlates exist, such as correlations of fluctuations among harmonic components other than vibrato, though we have not been able to discover them in the present flute sounds. Further investigations involving other flute sounds will yield useful information.

5. CONCLUSION

We investigated the relationship between amplitude and frequency fluctuations of harmonics and the perceived quality of flute tones with vibrato. Listener's preference for flute tones was found to be affected by the degree of intensification or attenuation of minute and irregular frequency and amplitude fluctuations above 13 Hz.

Also, we investigated the physical properties of the fluctuations that affect perceived quality of flute tones, by synthesizing fluctuation waves of harmonics. The results of evaluation by test subjects show that there was no perceived difference in quality between the original tone and synthesized tones with fluctuations that were synthesized by randomization of the phase spectra of the original fluctuations. In contrast, synthesized tones with fluctuations that were synthesized from filtered noise were perceived to be significantly inferior to the original tone. These results suggest that spectral variation of fluctuation waves which is at higher frequency and lower amplitude than spectral variation of vibrato influences perceived quality.

ACKNOWLEDGEMENTS

We would like to thank Prof. Iwasaki for recording flute sounds and taking part in the listening experiment. This research was supported in part by a grant from the collaboration program of Tokyo University of Information Sciences in 1998, 1999 and 2000.

REFERENCES

[1] T. Kunii and Y. Ando, "Physical properties of tones of the musical instrument in connection with its tone quality (III) (Harmonic component fluctuation observed in sustained notes of the flute)", *Proc. Autumn Meet. Acoust. Soc. Jpn.* 2-4-8, 329–330 (1973) (in Japanese).

[2] Y. Ando and T. Shima, "Physical properties of the flute tone and its subjective goodness of quality", *J. Jpn. Musicol. Soc.* **22**, 68–81 (1976).

[3] Y. Ando and H. Yamaya, "Physical correlates of the lowest note (Tsutsu-ne) of Shakuhachi", *Proc. Autumn Meet. Acoust. Soc. Jpn.* 3-3-13, 335–336 (1982) (in Japanese).

[4] K. Okao, "Physical correlates of double-reed instrument tones", *Master's thesis, Kyushu Institute of Design* (1979) (in Japanese).

[5] T. Ikemiya, "Physical correlates of horn tones", *Graduate thesis, Kyushu Institute of Design* (1979) (in Japanese).

[6] Y. Ando and T. Shima, "Physical properties of sustained part of the treble recorder tone and its subjective excellence of quality", *J. Jpn. Musicol. Soc.* **23**, 81–101 (1977).

[7] Y. Ando and K. Sakagami, "Non-periodic fluctuations of amplitude & frequency of each harmonics of treble recorder tones", *Proc. Autumn Meet. Acoust. Soc. Jpn.* 2-9-5, 719–720 (1994) (in Japanese).

[8] M. Kato, "On a method for accurate measuring of pitch frequency of musical tones", *Tech. Rep. Musical Acoust., Acoust. Soc. Jpn.* MA95-55, 1–8 (1996) (in Japanese).