800

Some empirical observations on sound level properties of recorded piano tones

Bruno H. Repp
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511-6695

(Received 3 April 1992; accepted for publication 13 October 1992)

Preliminary to an attempt at measuring the relative intensities of overlapping tones in acoustically recorded piano music, this study investigated whether the relative peak sound levels of recorded piano tones can be reliably inferred from the levels of their two lowest harmonics, measured in the spectrum near tone onset. Acoustic recordings of single tones were obtained from two computer-controlled mechanical pianos, one upright (Yamaha MX100A Disclavier) and one concert grand (Bösendorfer 290SE), at a range of pitches and hammer velocities. Electronic recordings from a digital piano (Roland RD250S), which were free of mechanical and sound transmission factors, were included for comparison. It was found that, on all three instruments, the levels of the lowest two harmonics (in dB) near tone onset generally increased linearly with the peak root-mean-square (rms) level (in dB) as hammer velocity was varied for any given pitch. The slope of this linear function was fairly constant across midrange pitches (C2 to C6) for the first harmonic (the fundamental), but increased with pitch for the second harmonic. However, there were two sources of unpredictable variability: On the two mechanical pianos, peak rms level varied considerably across pitches, even though the strings were struck at nominally equal hammer velocities; this was probably due to the combined effects of unevenness in hammer–string interaction, soundboard response, and room acoustics. Moreover, for different pitches at *equal* peak rms levels, the levels of the two lowest harmonics varied substantially, even on the electronic instrument. Because of this variability, the relative levels of the first or second harmonic near onset provide only very rough estimates of the relative peak levels of recorded piano tones.

PACS numbers: 43.75.Mn

INTRODUCTION

The present investigation was stimulated by a practical problem. In the course of studying the expressive microstructure of acoustically recorded piano performances (Repp, 1990, 1992), the question arose how to best estimate the relative intensities of individual tones. In piano music there are usually several tones present at the same time, initiated simultaneously or in succession. Their intensities are thus confounded in any overall measure derived from the acoustic waveform, such as the maximum amplitude or the root mean square (rms) amplitude within some time window. To separate the simultaneous sounds, spectral analysis is necessary. Ideally, one would like to have a technique that automatically separates the spectra of simultaneous complex tones. This is a difficult problem, however, which could not be tackled by this author for a variety of reasons. For the present purpose, therefore, a simpler solution was sought.

It was hypothesized that a reasonable estimate of the relative peak levels of individual tones might be obtained by measuring the relative levels of their fundamental frequencies (first harmonics) in Fourier spectra. The pitch and approximate time of onset of each tone were assumed to be known from the musical score and from acoustic waveform measurements, respectively. Unless the pitch distance between simultaneous tones is very small, their fundamental frequencies should appear as separate peaks in their combined spectrum. The heights of these peaks should be independent of each other, unless the pitch relationship is such that a higher harmonic of the lower tone coincides with the fundamental of the higher tone (i.e., if they are an octave apart, or a twelfth, and so on). These cases could be dealt with by applying an empirically or theoretically derived correction for additivity of harmonics. In most instances, however, independence of fundamental frequency peaks in the spectrum can be assumed.

Before the problem of simultaneous tones can be addressed, however, it must be determined whether the relative peak levels of *single* tones can be inferred reliably from the relative levels of their fundamentals. Two functions must be determined: one that relates these two quantities for any given pitch, and another that shows how that relationship changes with pitch. It is known from theoretical investigations and measurements of piano string vibration (e.g., Hall and Askenfelt, 1988) that the relative strength of the higher harmonics in the spectrum increases as hammer velocity (and hence peak sound level) increases for a given pitch. There is apparently no information in the literature, however, about whether the increase in the level of any individual harmonic with peak level is linear or nonlinear. It is also known that, as pitch increases, the fundamental becomes increasingly prominent in the spectrum of piano tones, thus accounting for a greater proportion of the overall intensity (see Benade, 1990). Again, however, the precise nature of this functional relationship is apparently not specified in the literature. Because the fundamental can become rather weak

 0001-4966/93/021136-09$06.00

low pitches, due to poor radiation by the soundboard (see Fletcher and Rossing, 1991), the present study not only avoided the extremes of the pitch range, but also examined the levels of both the fundamental (first harmonic) and of the second harmonic.

Although the theoretical apparatus may be available to make fairly precise predictions of piano string vibration, which is subject only to variation arising from mechanical factors, piano sounds recorded by a microphone are subject to variation from two additional sources: the radiation characteristics of the soundboard and effects of room acoustics dependent on the position of the microphone. These effects are extremely complex and difficult to predict (Benade, 1990). Commercial piano recordings thus present a rather messy situation from a scientific viewpoint: The specific characteristics of the instrument, the room acoustics, and the microphone placement(s) are all unknown, not to mention any additional sound processing by recording engineers or band limitations and distortion on older recordings. It is impossible to infer the "original" piano sound from such a recording, let alone the hammer velocities that produced the sound; the recorded sounds must be accepted as they come. (After all, they are what the listener hears.) The purpose of this study was to examine whether piano tones recorded under conditions not unlike those of a commercial studio recording would show systematic relationships among the intensity measures of interest, and to assess the magnitude of the unsystematic variation to be reckoned with.

Although the investigation could, in principle, have been conducted on sounds recorded from a conventional instrument played manually, advantage was taken of the availability of two different computer-controlled mechanical pianos on which MIDI or hammer velocity could be specified precisely, so that different strings could be excited with comparable forces. The inclusion of recordings from two different instruments was expected to lend some generality to the conclusions drawn, as well as to provide information about some specific characteristics of each of these instruments. This information might prove useful for investigators in music psychology who plan to use computer-controlled pianos for the generation of materials for perception experiments. For comparison, electronic recordings from an instrument producing synthetic piano sound were included as well. These recordings served as a control or baseline, for they were free of any variability due to mechanical factors and sound transmission (resonance and absorption) effects. Of course, their representativeness depended on the realism of the sound synthesis algorithm; at the very least, however, they provided some information about the output of this proprietary algorithm, which may also be useful to researchers contemplating to use this instrument.

A recent precedent for an empirical investigation of piano acoustics from a psychomusicological perspective is the study of Palmer and Brown (1991). They investigated the maximum amplitude of piano tones as a function of hammer velocity in recordings obtained from a computer-controlled Bösendorfer 290SE concert grand piano located at the Media Lab of the Massachusetts Institute of Technology. Maximum amplitude was measured from digitized waveforms of acoustically recorded tones at several different pitches. Palmer and Brown found that this quantity increased as a linear function of hammer velocity, for velocities between 0.5 and 4 m/s. The slope of the linear function was similar for most pitches tested, but some pitches showed a different slope—an irregularity that remained unexplained. Palmer and Brown further showed that the maximum amplitude of two simultaneous tones was linearly related to the sum of the maximum amplitudes of the individual tones, with a slope somewhat below 1. The rather basic information provided by this empirical study seemed not to be directly available in the literature on piano acoustics, which generally takes an experimental physics approach. The present investigation follows in the footsteps of Palmer and Brown in that it provides some empirical results that may be of interest not only to music acousticians but also to music psychologists using computer-controlled pianos.

I. METHODS

A. Instruments

The first instrument recorded from was a Roland RD250S digital piano at Haskins Laboratories. This instrument uses proprietary "structured adaptive" synthesis algorithms, which are said to faithfully recreate the sounds of several different types of piano, including their variations in timbre with pitch and sound level. "Piano 1," which indeed has a fairly realistic—though still recognizably artificial—piano sound, was used for the present recordings (with intermediate "brilliance" setting). The instrument was MIDI-controlled from an IBM-compatible microcomputer using the FORTE sequencing program.

The second instrument was a Yamaha MX100A Disclavier upright piano located at Yale University's Center for Studies in Music Technology. It is a traditional mechanical instrument equipped with optical sensors and solenoids that enable MIDI recording and playback. Control was via a Macintosh IIcx computer using the PERFORMER sequencing program. The translation from MIDI velocities (ranging from 0–127) to actual hammer velocities was not known for this instrument.

The third instrument was a Bösendorfer 290SE concert grand piano located in the Music School at The Ohio State University. Like the Yamaha Disclavier, it is a traditional piano equipped with optical sensors and solenoids. Unlike the Yamaha, however, it comes with customized hardware and software that specifies actual hammer velocities (in m/s). In this study, however, the piano was MIDI controlled from a Macintosh II computer using STUDIO VISION software. The mapping from actual hammer velocities to MIDI velocities had been carried out previously by technicians at The Ohio State University.[1]

B. Recording procedure

The same array of sounds was recorded from each instrument. To keep the study within bounds, extremes of pitch and loudness were avoided. The pitches recorded ranged from C2–C6 (C4 = middle C, about 262 Hz) in steps of three semitones: C2, E♭2, G♭2, A2, C3,..., C6. Each tone was played once at each of five MIDI velocities: 20, 40, 60, 80, and 100. Each tone lasted several hundreds of milliseconds (well beyond the peak amplitude), and comparable intervals of silence intervened between successive tones.[2] All recordings were made on high-quality cassette recorders with Dolby B noise reduction.

The output of the Roland was recorded electronically, going directly from the output jack to the tape recorder, to avoid any effects of sound transduction and room acoustics. The Yamaha upright was recorded with a Sennheiser MD409U3 microphone placed on a chair located about 1 m in front of the keyboard. The lid of the piano was open. The Bösendorfer grand piano was recorded with an AKG 451 cardioid (condenser) microphone placed on the right side, about 1 m from the open lid.[3]

C. Acoustic analysis

Each sequence of recorded tones was played back with Dolby B on a high-quality cassette deck and was digitized on a VAX 11/780 computer using the Haskins Laboratories Pulse Code Modulation System (Whalen *et al.*, 1990). The sampling rate was 10 kHz, with low-pass filtering at 4.9 kHz but without high-frequency pre-emphasis. It was assumed that frequencies above 5 kHz would contribute little to the peak level of the sounds.

The *peak root-mean-square (rms) level* in dB of each tone was measured by moving a 12.8-ms time window in 6.4-ms steps across each sampled data file, computing the rms level in each window, and picking the maximum value.[4] The resolution was 0.1 dB. Since lower pitched piano tones have a slower amplitude rise time than higher pitched tones, the location of the maximum moved closer to tone onset as pitch increased.

The *levels of the first and second harmonics (H1 and H2)* in dB near the onset of each tone were determined (to the nearest 0.5 dB) from a display of the Fourier spectrum of each tone. The spectrum was computed from a 51.2-ms time window whose left edge coincided with the onset of sound energy, as determined visually in magnified waveform displays.[5] Since the amplitude rise was partially or wholly included in this window, and since rise time decreased with pitch, it may have had an indirect effect on H1 and H2 level.

Since only a single token of each pitch-velocity combination had been recorded from each instrument, no estimate was available of the variability in peak rms, H1, and H2 levels for the same key struck repeatedly at the same MIDI velocity. This variability, however, was expected to be small and randomly distributed over all measurements. Hall and Askenfelt (1981) and Palmer and Brown (1991), who did obtain multiple tokens of each tone, commented on the small token variability. As will be seen below, the present functions relating sound level to pitch at different MIDI velocities were strongly parallel, suggesting little uncontrolled within-pitch variability.[6]

It should be emphasized that the absolute dB values shown in the figures below reflect the different recording and playback levels employed for each instrument, and hence are not interpretable. Only relative differences are of interest. Due to the different methods of computation, peak rms levels are only about half the magnitude of H1 and H2 levels; only the latter are directly comparable to each other, as they are derived from the same spectra.

II. RESULTS AND DISCUSSION

A. Level measures as a function of pitch and MIDI (hammer) velocity

1. Peak rms level

For a constant nominal MIDI velocity, which represents or simulates a constant hammer velocity, peak rms level was expected (perhaps naively) to be constant or slowly changing across different pitches. Figure 1 shows peak

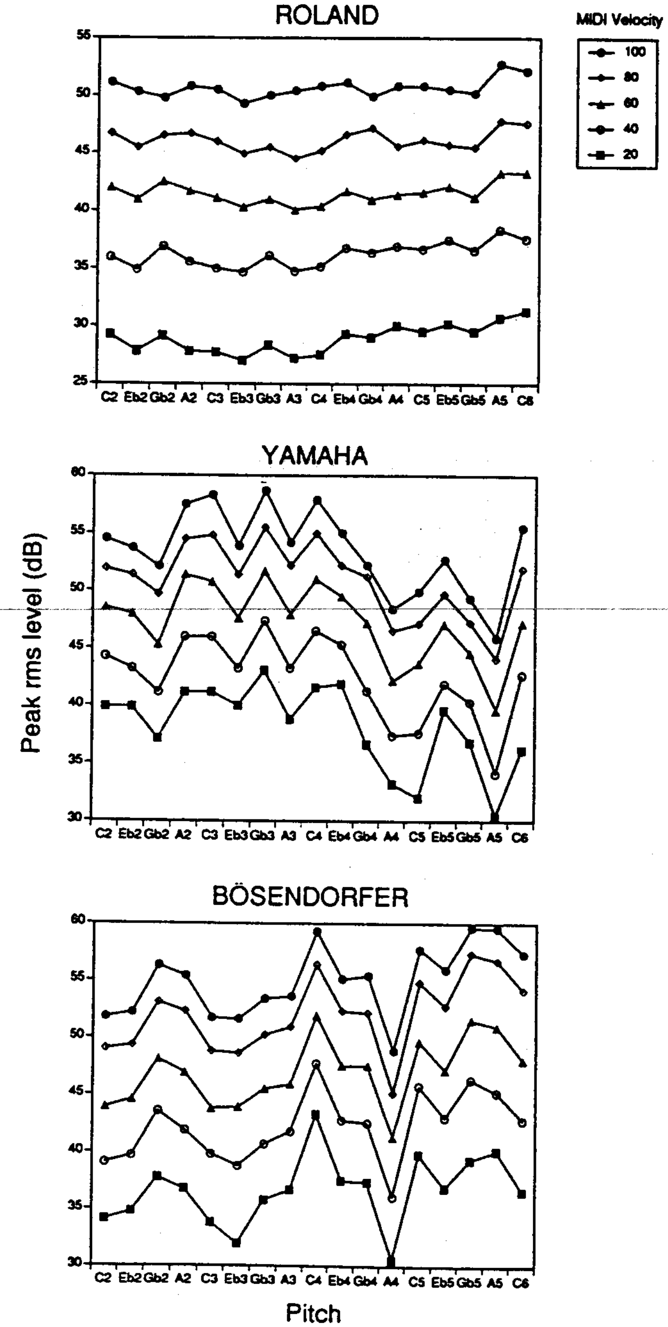

FIG. 1. Peak rms level as a function of pitch and MIDI velocity on three instruments.

rms level as a function of pitch at each of the five MIDI velocities, separately for each instrument. On the Roland, peak rms level was indeed reasonably constant, although at low MIDI velocities, there was a tendency for higher pitched tones to be a few dB more intense than lower pitched tones. The range of variation at any given MIDI velocity did not exceed 4 dB.

The Yamaha and Bösendorfer data were quite different. There were dramatic variations in peak rms level as a function of pitch. The range of variation at a constant nominal MIDI velocity was as much as 13 dB on each instrument. Moreover, the pattern of variation seemed unsystematic and unrelated between the two instruments. The curves for different MIDI velocities were closely parallel, however, indicating only a small contribution of measurement error or token variability. The observed variation may have been superimposed on a more systematic change in peak rms level with pitch, but any such underlying trend was difficult to discern because of the large variability. If anything, peak rms level tended to decrease with pitch on the Yamaha, while it tended to increase on the Bösendorfer.

In theory, this quasirandom variation in the relative levels of different piano tones could have at least four causes: (a) variation in the force with which the hammers were set into motion, due to imprecise calibration of the electronic components; (b) variation in the mechanical action and in the hammer–string interaction, due to differences in friction, hammer surface, area of hammer-string contact, etc.; (c) variation due to the resonant characteristics of the piano soundboard; and (d) variation due to acoustic frequency absorption or enhancement by surfaces and objects in the room. The relative contributions of these sources are unknown. Fletcher and Rossing (1991, p. 327) display a figure from a study by Lieber (1979) showing similar variation in the relative sound levels of (upright) piano tones when the strings were struck at a constant force. Since Lieber averaged over several different microphone positions and controlled physical force, his data reflect variation mainly due to (b) and (c).[7]

Figure 1 shows that, orthogonal to the variation in relative peak level across pitches, there was a systematic increase in peak rms level with MIDI velocity on each instrument—certainly an expected result. Moreover, the nature of that increase seemed similar at all pitches; hence, the parallelism of the functions in each panel of Fig. 1. It made sense, therefore, to average across all pitches on each instrument to examine the relationship between *average* peak rms level and MIDI velocity. In each case, this function was mildly nonlinear and negatively accelerated; it was fit well by a quadratic (second-order polynomial) curve. In other words, peak rms level tended to increase faster at low than at high MIDI velocities. The function was steepest and most curved for the Roland and flattest and least curved for the Yamaha, due to differences in dynamic range. The dynamic range (the difference between the average peak rms values at MIDI velocities of 20 and 100) was 23 dB on the Roland, 15 dB on the Yamaha, and 18 dB on the Bösendorfer.[8]

Since the nominal hammer velocities (in m/s) were known on the Bösendorfer, the function relating average peak rms level and hammer velocity could be examined. It was approximately logarithmic. This is in agreement with the results of Palmer and Brown (1991) who found maximum amplitude (measured on a linear scale, not in dB) to be a linear function of hammer velocity.[9]

2. H1 level

It is known that the relative prominence of the fundamental frequency in the spectrum of piano tones increases

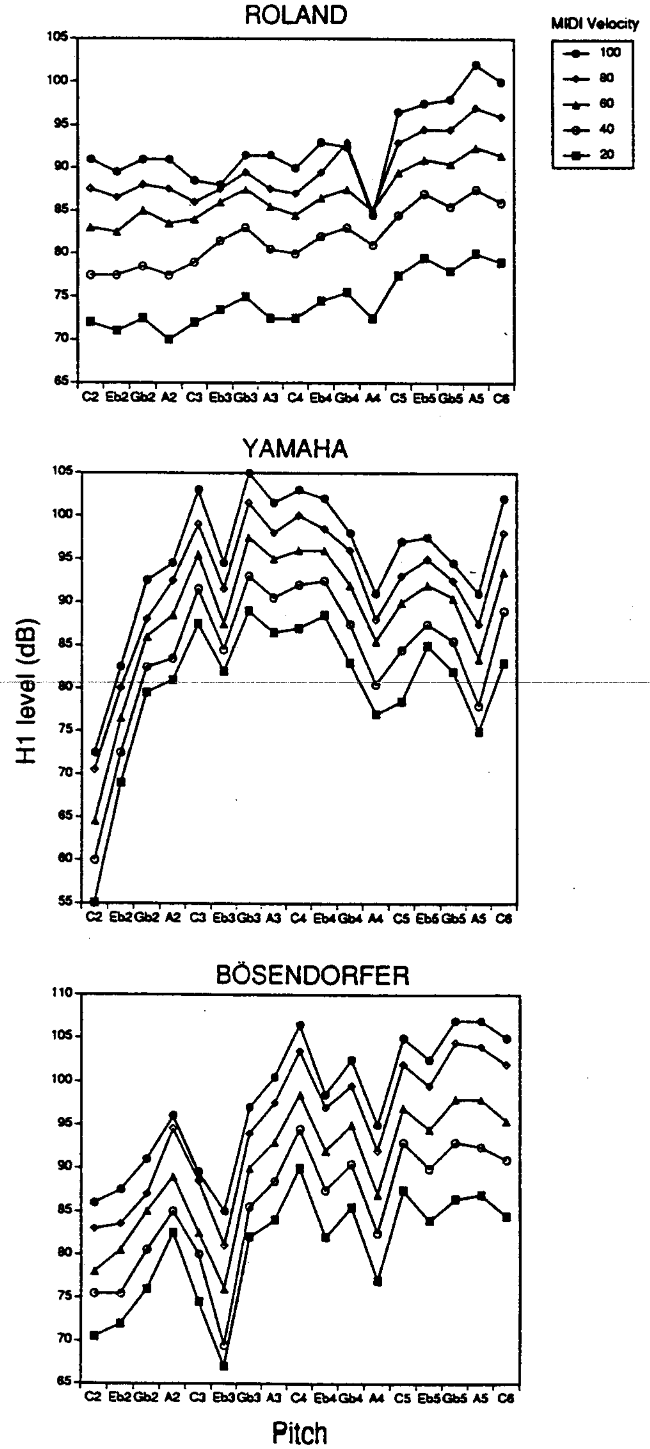

FIG. 2. H1 level as a function of pitch and MIDI velocity on three instruments.

with pitch (Hall and Askenfelt, 1988). Therefore, the function relating H1 level to pitch at a constant MIDI velocity was expected to be rising rather than flat. However, given the unsystematic variation in peak rms level across pitches on the two mechanical pianos (Fig. 1), comparable irregularities in the relationship of H1 level to pitch were to be expected for these instruments.

The relevant data are shown in Fig. 2. The Roland results indeed show gradually rising functions, with variability comparable in magnitude to that of peak rms level. However, there were a few unexpected (and unexplained) failures of H1 level to increase further at high MIDI velocities (at pitches E♭3, G♭4, and especially A4). Apart from these flukes, however, the functions for different MIDI velocities were parallel. It appears that the Roland synthesis algorithm successfully simulates the expected increase in relative H1 level with pitch.

The Yamaha data were quite different. At pitches below G♭2, there was a precipitous decline in H1 level, evidently due to the "critical frequency" of the soundboard, below which the sound is not radiated effectively (see Fletcher and Rossing, 1991, p. 326). Above G♭2, there was no systematic increase in H1 level with pitch. Instead, there was major variability, similar in magnitude to that observed for peak rms level. (Note that the ordinate scale is compressed in Fig. 2 relative to Fig. 1.)

The Bösendorfer data did show an overall trend for H1 level to increase with pitch, but with major irregularities, as expected. The different MIDI velocity functions were fairly parallel. The change of H1 level with pitch seemed quite different on the Bösendorfer than on the Yamaha, although on both instruments there were dips in level at pitches E♭3 and A4. This may be a coincidence, however. There was no abrupt fall-off in H1 level at the lowest frequencies, presumably due to the lower critical frequency of the larger soundboard of a grand piano (cf. Suzuki, 1986).

Since the functions in Fig. 2 were generally quite parallel, except for a few anomalies in the Roland data, it seemed justified to examine the relationship between average H1 level (averaged across pitches) and MIDI velocity. As for peak rms level, these functions were negatively accelerated (quadratic), more so on the Roland than on the other two instruments. The average dynamic ranges of H1 level are 18 dB on the Roland, 15 dB on the Yamaha, and 17 dB on the Bösendorfer. These values are similar to those for peak rms level, except on the Roland, where the range of H1 level is more restricted than that of peak rms level.

On the Bösendorfer, average H1 level could be examined as a function of nominal hammer velocity. This function was logarithmic, similar to that for peak rms level.

3. H2 level

It has been observed previously that, whereas the relative prominence of H1 in the piano tone spectrum increases with pitch, that of H2 decreases (Hall and Askenfelt, 1988). Therefore, H2 level was expected to decrease as a function of pitch at each MIDI velocity. Again, however, unsystematic variability was to be expected on the mechanical pianos. Figure 3 shows the H2 data.

On the Roland, random variability across pitches was fairly small, and the predicted decrease in H2 level with increasing pitch was present. However, this decrease was much more pronounced at lower MIDI velocities, leading to a fanning out of the functions in the figure. It is well known that higher hammer velocities increase the relative prominence of higher harmonics in the spectrum and thereby brighten the timbre of the piano sound (Benade, 1990). At lower pitches, this effect is probably conveyed primarily by harmonics above H2, whereas in higher pitches, which have fewer harmonics, it shows up increasingly in H2. At least, this is how it was simulated in the Roland synthesis algorithm.

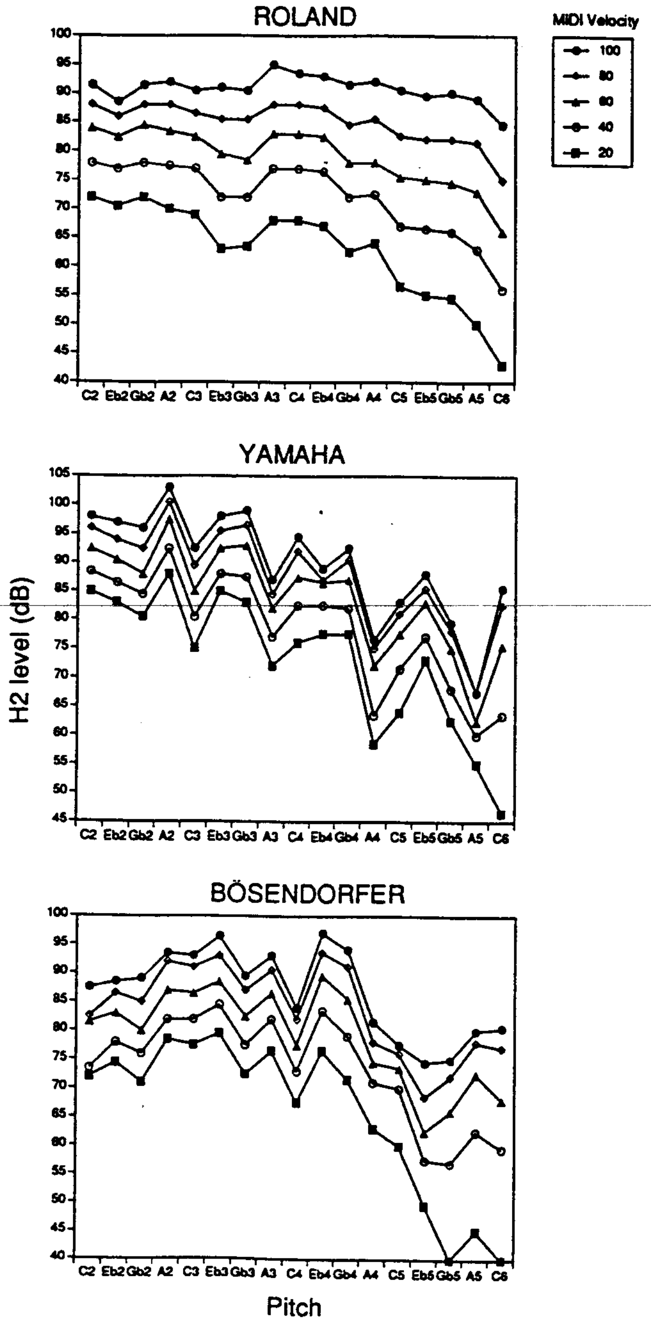

FIG. 3. H2 level as a function of pitch and MIDI velocity on three instruments.

On the Yamaha, the overall decline in H2 level was more pronounced but marred by the expected unsystematic variability. (That there were dips at A3, A4, and A5 may well be a coincidence.) The decline was present at all MIDI velocities, and there was no significant fanning out of the functions, except at the highest pitch (C6). On the Bösendorfer, there was little overall decline in H2 level up to G♭4; a precipitous decline between G♭4 and E♭5 was followed by another plateau. There was a pronounced fanning out of the functions in the highest octave, mainly due to H2 level becoming very weak at the lowest MIDI velocity.

The functional relationship of H2 level to MIDI velocity, like that of H1 level to MIDI velocity, was nonlinear and negatively accelerated (quadratic) on all three instruments. Because of the fanning out observed in Fig. 3, this function became steeper at higher pitches. On the Bösendorfer, H2 level could also be examined as a function of nominal hammer velocity. That relationship was approximately logarithmic and quite similar to that for H1, though the logarithmic fit was less good for H2.

B. Relationships among level measures

The principal aim of this study was to examine whether peak rms level could be estimated from H1 and/or H2 level (and vice versa). Ignoring, at first, the unsystematic variation in relative level across pitches, H1 and H2 level were examined as a function of average peak rms level at each individual pitch. In most instances, these functions were strikingly linear. Significant deviations from linearity (defined arbitrarily, but rigorously, as r-squared <0.99) were noted. Figure 4 plots the slopes of these linear functions as a function of pitch; deviations from linearity are indicated by asterisks.

For H1 level, the slopes on each instrument varied mainly between 0.8 and 1, with a tendency to increase with pitch. On the Roland, 6 of the 17 pitches examined showed a deviation from linearity in the direction of negative acceleration and hence had a reduced linear slope; the remaining 11 pitches yielded linear functions. On the Yamaha and Bösendorfer, all functions were linear. The lowest pitch (C2) on the Yamaha, for which H1 level was unusually low, showed a steeper slope than other pitches; otherwise, the results were comparable for the two mechanical instruments.

The functions relating H2 level to peak rms level within pitches also were predominantly linear. On the Roland, the slope of the linear function increased steadily with pitch, from below 1 (similar to H1) to almost 2. The slopes were abnormally steep for pitches E♭3 and G♭3, but there was no deviation from linearity. On the Yamaha, the H2 slopes were more variable and did not increase systematically with pitch until about A4. The slope was abnormally low for A5 and abnormally high for C6. Significant nonlinearities were observed at two pitches, E♭4 and C6. On the Bösendorfer, the slopes increased more gradually with pitch but showed significant nonlinearities in the highest octave. Thus the function relating H2 level and peak rms level, unlike the corresponding function for H1, was not always linear on the mechanical instruments and also more variable in slope. On the Roland, on the other hand, H1 was a less consistently linear function of peak rms level than was H2.

The linear relationship between H1 level and peak rms level within pitches, and the relative stability of that relationship across pitches on the mechanical instruments, are encouraging for the purpose of estimating one from the other. However, a crucial question remains. It was observed that both peak rms level and H1 level varied substantially and unsystematically as a function of pitch (Figs. 1 and 2). If

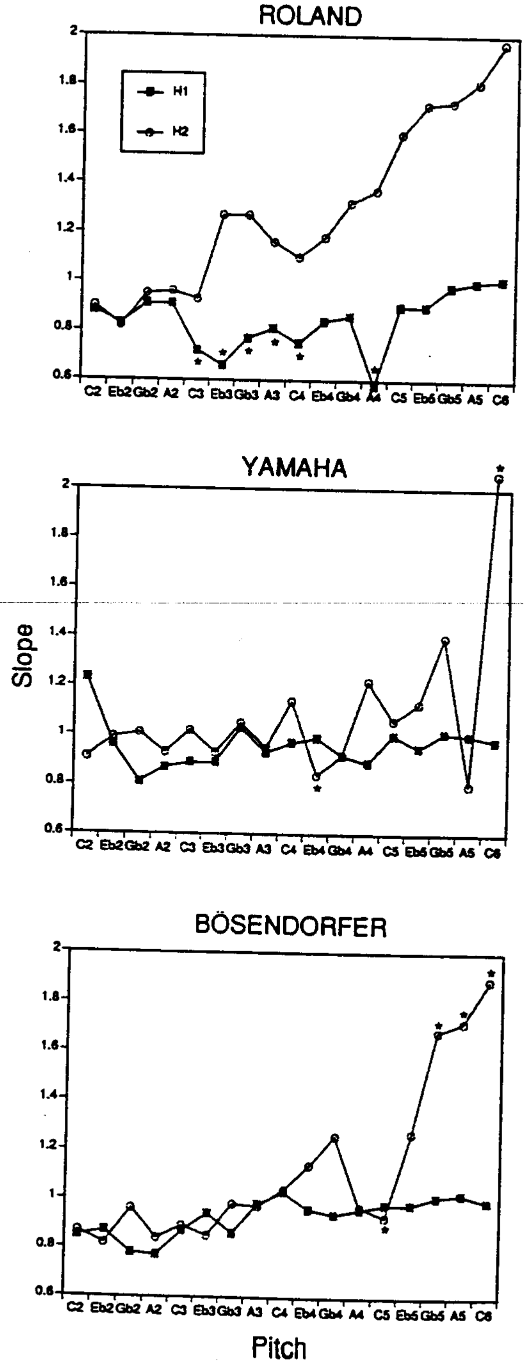

FIG. 4. Slopes of the linear functions relating H1 or H2 level to peak rms level as a function of pitch on three instruments. Asterisks indicate deviations from linearity (r-squared <0.99).

these variations were highly correlated, so that H1 level went up and down with quasirandom variations in peak rms level, then the variability would not prevent estimation of one from the other. However, if these variations were only weakly correlated, the predictability of peak rms level from H1 level (or vice versa) would be seriously impaired.

This issue is addressed partially in Fig. 5, which plots the relationship between H1 level and peak rms level for a constant representative MIDI velocity, arbitrarily selected to be 60. The data points in these scatter plots represent the different pitches. On the Roland, there was a very narrow range of peak rms levels (corresponding to the variability across pitches in Fig. 1) but a much wider range of H1 levels: The former varied by less than 4 dB, while the latter varied over a 10-dB range. There was only a weak correlation between the two; *for the same peak rms level, H1 level could vary by as much as 8 dB.* Also, note that the slope of a linear function relating the two variables would be much steeper than the slope of the function relating these variables when pitch was fixed and MIDI velocity varied (Fig. 4).

On the Yamaha, the relationship was even poorer, despite a much wider range of peak rms levels. This was mainly due to the exceptionally low H1 levels at the two lowest pitches (cf. Fig. 2). Without these two data points, the correlation would clearly be higher, and the slope of the linear regression function would be close to 1. However, as on the Roland, H1 level would still vary by as much as 8 dB for the same peak rms level.

A correlation between H1 level and peak rms level was also observed on the Bösendorfer, but here, as on the Roland, the slope of a linear regression function would be about twice as steep as that of the function for fixed pitch (Fig. 4). H1 level varied by 10 dB or more for tones of similar peak rms level.

When attempting to infer peak rms level from H1 level measured in acoustic recordings of an unknown instrument, within-pitch and between-pitch variability cannot be distinguished. Thus what is most relevant in the present data is the overall correlation between the two acoustic variables, across all pitches and MIDI velocities. These correlations are not as high as one would like them to be: 0.90 on the Roland, but only 0.61 on the Yamaha (where the two lowest pitches provided highly deviant data), and 0.84 on the Bösendorfer. Judging from scatter plots of all the data points, the error in estimating peak level from H1 level can exceed ± 5 dB.

Even larger variability was observed in H2 level for tones with identical or highly similar peak rms values. The correlation between H2 level and peak rms level across all pitches and MIDI velocities was 0.83 on the Roland, and the range of peak rms values for a given H2 level could be as wide as 15 dB at the lower levels. A similar degree of variability was observed on the Yamaha, where the correlation was 0.80. Finally, on the Bösendorfer, the correlation was dismal, only 0.44, and estimation errors in excess of ± 10 dB would seem possible.

A final possibility to be explored was that H1 and H2 levels combined might show a tighter relationship to peak rms level than either of these variables alone, because of a trade-off between them. One method was to simply average the H1 and H2 levels. On the Roland and the Yamaha, this led indeed to an improvement in the correlation with peak rms level over the correlations for H1 and H2 individually. On the Roland, the correlation was 0.96, and errors in estimating peak rms level would generally be within ± 2 dB. On the Yamaha, the correlation was 0.90, and estimation errors could be as large as ± 5 dB. However, on the Bösendorfer the correlation was only 0.82, not higher than that for H1 alone, and estimation errors in excess of ± 5 dB seemed possible. Alternatively, when the relation between H1 and H2 levels and rms peak level was examined by means of multiple regression, the multiple correlations were 0.97, 0.91, and 0.91, respectively, on the three instruments. The Bösendorfer data did look somewhat better from that perspective, but the estimation error to be reckoned with was still substantial.

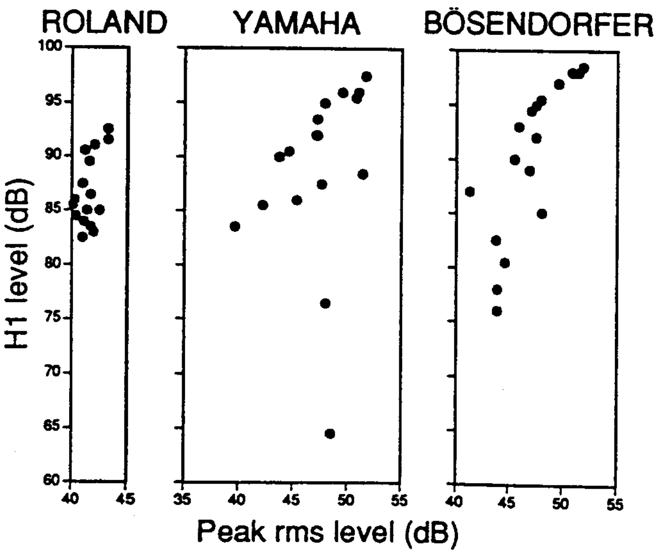

FIG. 5. H1 level as a function of peak rms level at a constant MIDI velocity of 60, on three instruments.

III. SUMMARY AND CONCLUSIONS

This empirical study was conducted to investigate how accurately the peak rms level of recorded piano tones could be estimated from the levels of their first two harmonics, measured near onset. Within pitches, where the level variation was caused almost entirely by variations in hammer velocity, the relationship between peak rms level and H1 level was linear on both mechanical instruments investigated, with H1 level increasing at or slightly below the rate of peak level increase. Between C2 and C4, the relationship between H2 level and peak rms level was similar to that between H1 level and peak rms level; above C4, however, H2 level tended to increase faster than peak level, and large variations in the slope of the linear function, as well as some significant nonlinearities, were found on both mechanical instruments. These results suggest that H1 level is a more reliable predictor of peak rms level than is H2 level, although a combined measure of H1 and H2 may be superior to H1 alone.

Although the relationship between H1 level and peak rms level was quite regular within pitches, there were two unexpectedly large sources of variability across pitches. One of them concerned peak rms level, the other the relationship between H1 and H2 levels and peak level. The first kind of variation was observed primarily on the two mechanical pianos; it was rather small on the electronic instrument. Its most likely cause lies in the resonance patterns of the piano sound board, though irregularities of the mechanical action from key to key and room acoustics may also play a role. The variation was unsystematic (though not totally random, it seems) and showed quite different patterns on the Yamaha and Bösendorfer pianos. The stability of the pattern for any given instrument across variations in microphone placement and room acoustics was not investigated here; the study merely established the existence and magnitude of such variability in one essentially arbitrary recording situation. Although the variability may pattern differently if the recording situation were changed, its magnitude would probably be similar.

The other kind of unsystematic variability, that of H1 and H2 levels across tones of different pitch having comparable peak rms levels, was observed not only on the mechanical pianos but on the electronic piano as well; this suggest that its origin lies neither in the hammer–string interaction, nor in the soundboard response, nor in room acoustics, which creates a puzzle. However, it is possible that the Roland's proprietary synthesis algorithm models piano sounds exhibiting natural spectral variation due to some of these sources, while disregarding the more gross differences in overall sound level. The cause of this spectral variation is not well understood at present, nor is its stability known for any given instrument.[10] Again, the present study merely demonstrates the variability; an explanation would require a more systematic inquiry.

The first type of variation, that in peak level with pitch, is worrisome, but perhaps more to the pianist or listener than to the scientist analyzing the recorded sound. Pianists may have to adjust to it and compensate for it to some extent in order to realize the fine dynamic gradations that musical expression demands.[11] However, it is impossible in any case to infer the original hammer velocity from the recorded sound because of additional distortions introduced by microphone placement, room acoustics, and sound engineering. The performance analyst must take the recorded sound at face value, for what it is worth. This first kind of variation, therefore, is not really an obstacle to estimating peak rms level from H1 (and H2) level.

The second kind of variation, however, significantly reduces the accuracy of such estimates. As Fig. 5 demonstrated, the error may be considerable. The present data are not sufficient to derive a precise estimate of the average error, but errors as large as ± 5 dB seem possible. The average error is going to be smaller, of course, and the possibility of obtaining sufficiently systematic and interpretable results in the analysis of recorded piano performances is by no means ruled out.

ACKNOWLEDGMENTS

This research was made possible through the generosity of Haskins Laboratories (Michael Studdert-Kennedy, president) and was supported in part by NIH (BRSG) Grant RR-05596. I am grateful to Jonathan Berger and Jack Vees (Center for Music Technology, Yale University) for permitting and helping me to record their Yamaha Disclavier, and to Allan Dudek and Pete Tender (School of Music, The Ohio State University) for providing the Bösendorfer recording, as well as further information and advice. For helpful comments on an earlier version of the manuscript, I am indebted to Donald Hall, Caroline Palmer, and an anonymous reviewer.

[1] This had been done by matching the peak sound levels (measured with SOUND DESIGNER software) of Bösendorfer tones as closely as possible to those of MIDI-controlled "Piano 1" tones produced at the same pitch on a Roland RD1000 digital piano. If the sound synthesis algorithms of the Roland RD250S and RD1000 digital pianos are identical (which the author believes to be the case), then the MIDI scale of the Bösendorfer should correspond to that of the Roland RD250S. The present measurements indeed showed them to be quite similar, though not identical. The relationship between MIDI velocity (MV) and nominal hammer velocity (HV) in m/s on the Bösendorfer was approximately logarithmic: $MV = 51.7 + 45.7 \ln(HV)$.

[2] The exact durations, which varied across instruments, are irrelevant because the analyses concerned only the initial portions (less than 100 ms) of the tones.

[3] No attempts were made to vary the recording environment or assess the effects of room acoustics. The relative arbitrariness of each recording situation corresponds to the equally arbitrary (or at least unknown) circumstances of commercial recordings.

[4] This method differs from that of Palmer and Brown (1991), who determined the maximum amplitude on a linear scale from a waveform display. The peak rms level measure used here is approximately a logarithmic transform of their measure. Note that absolute values have no meaning in either case, since they depend on recording and playback levels.

[5] This window was not the same as that from which the peak rms level was derived, although for the higher pitched tones it overlapped or included it. There were two practical reasons for using different windows. One was that, in his ongoing performance analysis project, the author did not have the resources to compute multiple spectra for each of thousands of tones, which would have been required to determine peak H1 and H2 levels; measurements of H1 levels from spectra at tone onsets constitute the data still awaiting analysis. The second reason was that the ILS software used to compute the Fourier spectrum did not yield an estimate of rms level, whereas the in-house program used to determine rms level did not permit the computation of a spectrum with a predetermined time window. With some further programming effort, it would, of course, have been possible to compute the rms level over the initial 51.2 ms of each tone and/or the H1 and H2 levels in the 12.8-ms window from which the peak rms level was computed. These additional measures, however, would only have provided intermediate steps in the concrete problem of making sense of the H1 level data obtained near tone onset.

[6] Another relevant piece of evidence comes from the Yamaha recordings, which originally included each tone played with a MIDI velocity of 120. These sounds, however, proved to be nearly identical to those with a MIDI velocity of 100, suggesting that the range beyond 100 was not functional on this instrument. From C2 to G♭3, the peak rms levels of tones with MIDI velocities of 100 and 120 differed by no more than ± 0.2 dB. From A3 to C6, the tones with the nominally higher MIDI velocities were from 0.1 to 0.9 dB more intense, suggesting some marginal effectiveness of velocity specifications beyond 100. The random variability of peak rms level due to inherent unreliability (token variation) and measurement error thus was surely well below 0.5 dB.

[7] The author has not seen Lieber's original study. According to Rossing (personal communication), the spacing of ticks on the (unlabeled) *y* axis in the reproduced figure is 10 dB, which makes the variation depicted similar in magnitude to that observed here. See also Savage *et al.* (1991, Fig. 6) for some harpsichord data showing again comparable variation.

[8] That the Roland had a larger dynamic range than the real pianos was unexpected, but perhaps this was an artifact of the electronic recording procedure, which avoided any loss of energy due to sound transmission and absorption. Although less than two-thirds of the total MIDI velocity range (0–127) was used, there was evidence that the Yamaha was not differentially responsive at the upper end of the MIDI scale (see footnote 6); a similar observation was made informally at the lower end of the scale. Palmer and Brown (1991) found that their Bösendorfer (a different specimen from the one recorded here) did not respond differentially to low hammer velocities. Thus the dynamic ranges found here may not be far from the maximum capacity of the mechanical instruments, at least under MIDI control.

[9] Their Fig. 3, however, suggests a poor fit for some pitches. A general problem with their figures is that they do not identify the data points for individual pitches, which makes it difficult to gauge the consistency of their data. Palmer and Brown also found a deviant (faster or slower, but still linear) growth in maximum amplitude with hammer velocity for some pitches. In the present data, however, the variability of the function relating peak rms level to MIDI or hammer velocity across individual pitches seemed relatively small. On the other hand, the large variability in relative peak level across pitches observed here was less evident in the Palmer and Brown study. Apart from the fact that they used a smaller number of pitches, the reasons for these differences in findings are not clear at present.

[10] Pilot data had been collected both on the Roland and the Yamaha for the pitches C3, C4, and C5. On both instruments, H1 level for C5 turned out to be at least 5 dB above the H1 levels for C3 and C4. This corresponds to the present data for the Roland (cf. Fig. 2), but not at all to those for the Yamaha, which show just the opposite. The pilot recordings from the Yamaha had been obtained with the microphone positioned inside the piano. The pattern of the observed variation thus may depend (at least) on microphone placement.

[11] Note, however, Benade's (1990) comment that "...the ordinary sound pressure recipes even for adjacent piano notes look unrecognizably different from one another when measured, although they may sound very well matched to our ears" (p. 437). The level measures examined here are acoustic, not perceptual primitives, and it is quite possible that listeners do not perceive the measured variability as such.

Benade, A. H. (1990). *Fundamentals of Musical Acoustics* (Dover, New York) (reprint of 1976 edition by Oxford University Press).

Fletcher, N. H., and Rossing, T. D. (1991). *The Physics of Musical Instruments* (Springer-Verlag, New York).

Hall, D. E., and Askenfelt, A. (1988). "Piano string excitation V: Spectra for real hammers and strings," J. Acoust. Soc. Am. **83**, 1627–1638.

Palmer, C., and Brown, J. C. (1991). "Investigations in the amplitude of sounded piano tones," J. Acoust. Soc. Am. **90**, 60–66.

Repp, B. H. (1990). "Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists," J. Acoust. Soc. Am. **88**, 622–641.

Repp, B. H. (1992). "Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's 'Träumerei,' " J. Acoust. Soc. Am. **92**, 2546–2568.

Savage, W. R., Kottick, E. L., Hendrickson, T. J., and Marshall, K. D. (1991). "Air and structural modes of a harpsichord," J. Acoust. Soc. Am. **91**, 2180–2189.

Suzuki, H. (1986). "Vibration and sound radiation of a piano soundboard," J. Acoust. Soc. Am. **80**, 1573–1582.

Whalen, D. H., Wiley, E. R., Rubin, P. E., and Cooper, F. S. (1990). "The Haskins Laboratories' pulse code modulation (PCM) system," Behav. Res. Methods Instrum. Comput. **22**, 550–559.