

Effects of harmonics on relative pitch discrimination in a musical context

LAUREL J. TRAINOR

McMaster University, Hamilton, Ontario, Canada

The contribution of different harmonics to pitch salience in a musical context was examined by requiring subjects to discriminate a small ($\frac{1}{4}$ semitone) pitch change in one note of a melody that repeated in transposition. In Experiment 1, performance was superior when more harmonics were present (first five vs. fundamental alone) and when the second harmonic (of tones consisting of the first two harmonics) was in tune compared with when it was out of tune. In Experiment 2, the effects of harmonics 6 and 8, which stand in octave-equivalent simple ratios to the fundamental (2:3 and 1:2, respectively) were compared with harmonics 5 and 7, which stand in more complex ratios (4:5 and 4:7, respectively). When the harmonics fused into a single percept (tones consisting of harmonics 1, 2, and one of 5, 6, 7, or 8), performance was higher when harmonics 6 or 8 were present than when harmonics 5 or 7 were present. When the harmonics did not fuse into a single percept (tones consisting of the fundamental and one of 5, 6, 7, or 8), there was no effect of ratio simplicity.

This paper examines the contribution of different harmonics to relative pitch discrimination in a musical context. *Relative pitch discrimination* refers to the ability to compare the pitch interval (i.e., distance on a log frequency scale) between one set of two tones and another, where the fundamentals of the tones are at different absolute frequencies (e.g., the distance between tones of 100 and 150 Hz is equivalent to the distance between tones of 200 and 300 Hz, as the ratio between them is 2:3 in both cases). Pitch discrimination in isolation has been shown to be better for complex over sine tones with relatively low fundamental frequencies (see, e.g., Fastl & Weinberger, 1981; Platt & Racine, 1985). The first question addressed in the present paper is whether this result generalizes to pitch discrimination in a musical context, as defined by a simple melody conforming to Western musical structure. The second question is whether certain harmonics, that is, certain ratio relationships to the fundamental, facilitate pitch perception more than others in a musical context.

Several researchers have indicated that some harmonics may be more important than others for facilitating pitch discrimination. The *dominance region* refers to the band of harmonics whose frequencies largely determine the perceived pitch. Different researchers have proposed somewhat different dominance regions. Ritsma (1967) found that the third, fourth, and fifth harmonics dominated the perception of pitch for fundamental frequencies between 100 and 400 Hz. Plomp (1967) found that the dominance re-

gion depended to some extent on the fundamental frequency: The fourth and higher harmonics dominated for fundamental frequencies up to 350 Hz, the third and higher for fundamentals between 350 and 700 Hz, the second and higher for fundamental frequencies between 700 and 1400 Hz, and the first for frequencies above 1400 Hz. Terhardt proposed a fixed-frequency dominance region centered around 700 Hz (Terhardt, Stoll, & Seewann, 1982).

Moore, Glasberg, and Peters (1985) attempted to determine the dominance of individual partials in determining the pitch of a complex by measuring the pitch shifts that accompanied the mistuning of each harmonic. Measurable pitch shifts were found only when one of the six lowest harmonics was changed. Further, there were large individual differences in that the 3 subjects showed maximal pitch shifts for different mistuned harmonics. However, graphs for individual subjects (Moore et al., 1985, Figures 2, 3, and 4) indicated that when one harmonic gave a *much* larger pitch shift than the others, it was usually either Harmonic 1 (the fundamental), Harmonic 2 (one octave above the fundamental), or Harmonic 4 (two octaves above the fundamental).

There is further evidence that octave harmonics may be especially important in pitch processing. In an interval discrimination task using two-tone complexes with missing fundamentals, where the tones were 3 or 4 harmonics apart, Houtsma (1979) found that performance depended greatly on the harmonic number of the lowest tone, a dependence that he claimed is not reflected in either the optimum processor (Gerson & Goldstein, 1978; Goldstein, 1973), virtual pitch (Terhardt, 1974; Terhardt et al., 1982), or pattern transformation (Wightman, 1973) models. Inspection of the data indicated that peaks in performance occurred when at least one of the two tones stood in an octave relationship to the missing fundamental.

The question of whether harmonics standing in certain ratio relationships to the fundamental facilitate pitch per-

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Science and Engineering Research Board of McMaster University. I am grateful to John Platt and three reviewers for helpful comments on an earlier draft. Correspondence should be addressed to L. J. Trainor, Department of Psychology, McMaster University, Hamilton, ON, Canada L8S 4K1 (e-mail: ljtr@mcmaster.ca).

ception more than other harmonics has not been systematically addressed. However, effects of frequency ratio simplicity have been studied in a number of other contexts. Pythagoras is credited with having discovered that the sounds produced by strings divided into simple ratios (e.g., 1:2; 2:3) sound more pleasing than those of strings divided into more complex ratios. Much more recently, Schellenberg and Trehub (1994b) reanalyzed the results of many pitch interval studies to show that the relative simplicity of frequency ratio between pairs of tones accounted for much of the variance in similarity judgments, consonance and dissonance judgments, and the relative ease of discriminating tone patterns. They used the reciprocal of the natural logarithm of the sum of the two integers forming the ratio, when the ratio is in its simplest form, as an index of ratio simplicity. The simplest ratio between the frequencies of two tones (other than the unison, 1:1) is 1:2, which results in the interval of an octave. This interval has a number of interesting properties. Across virtually all cultures, tones an octave apart are perceived as being similar (Burns & Ward, 1982; Dowling & Harwood, 1986). For example, in Western musical structure as well as Indian *rag* structure, notes an octave apart have the same *tone chroma*, that is, are given the same note name. The perceptual equivalence of tones an octave apart is also found in young infants, raising the possibility that octave equivalence reflects basic auditory functioning, independent of experience (Demany & Armand, 1984).

The second most simple ratio between the frequencies of two tones within the octave range is 2:3, which results in the interval of a perfect fifth. Perfect fifths are also common across musical systems (Kolinski, 1967; Nettle, 1956, p. 54; Roederer, 1979, pp. 146–147), although they are not as universal as octaves (see, e.g., Burns & Ward, 1982). A few researchers have examined the role of perfect fifth intervals in sequential processing (i.e., perfect fifths between successive tones). Melodies with perfect fifth intervals between prominent tones are easier to process than are those with more complex frequency relations. For example, Trainor (1991) found that adults were better able to detect a downward semitone (i.e., $\frac{1}{2}$ th octave) change in the third note of a Western-based melody ($C_4-E_4-G_4-E_4-C_4$) containing an approximate 2:3 ratio that resulted in a diminished triad containing a 32:45 ratio than they were able to detect the reverse. The effect of the simple frequency ratio was found to be independent of whether the melody was Western in structure or not. A non-Western melody containing a 2:3 ratio was found to be easier to distinguish from a similar melody containing a 32:45 ratio than the reverse (Trainor & Trehub, 1993a). Further, Western melodies transposed by intervals of 2:3 were easier to compare than were melodies transposed by other, more complex, intervals (Trainor & Trehub, 1993b). Schellenberg and Trehub (1994a) also found similar asymmetries in discrimination in more musically impoverished contexts. Divenyi and Hirsh (1974) found that temporal order discrimination of three-tone sequences was superior when the tones were related by small-integer frequency ratios than otherwise. Using an adjustment methodology, J. Elliot, Platt, and Racine (1987)

found that the accuracy of musical interval perception was directly related to the simplicity of the frequency ratio of the component notes. The findings of enhanced pitch processing for melodies with prominent perfect fifth intervals also extend to infants (Cohen, Thorpe, & Trehub, 1987; Trainor, 1991; Trainor & Trehub, 1993a, 1993b), suggesting that it may either be innate or involve precocious or innately guided learning (Gould & Marler, 1987).

Frequency ratio effects are also evident with simultaneous tones. Typically, studies have examined the perceived consonance (i.e., pleasantness) of tone combinations. When two sine wave tones are combined there is little effect of frequency ratio simplicity; rather, intervals smaller than a critical bandwidth are perceived as relatively dissonant (unless they approximate the unison), while intervals larger than a critical band are perceived as relatively consonant (Kameoka & Kuriyagawa, 1969a; Levelt, Van de Geer, & Plomp, 1966; Plomp & Levelt, 1965; Terhardt, 1974). This phenomenon is referred to as *tonal consonance*. Simultaneous complex tones whose fundamental frequencies stand in small-integer ratios are perceived as more consonant than those whose fundamental frequencies stand in more complex ratios (Ayres, Aeschbach, & Walker, 1980; Kameoka & Kuriyagawa, 1969b; Levelt et al., 1966; Plomp & Levelt, 1965; Terhardt, 1974; Vos, 1986). The usual explanation for this effect involves interactions between the harmonics. The simpler the ratio relation between the fundamental frequencies of the complex tones, the more harmonics they have in common. Thus, complex ratio relations between two simultaneous complex tones result in harmonics less than a critical bandwidth apart, which interact in a similar fashion to the sine wave tones described above, resulting in the perception of dissonance. The hypothesis regarding the relation between consonance and critical bandwidth has been further strengthened by Greenwood (1991), who showed that across frequency, the smallest separation between two tones that results in the perception of consonance corresponds to a constant distance along the basilar membrane.

Terhardt (1974, 1984) proposed that processing advantages for simple frequency ratios could result from exposure to the harmonic series that occurs in common complex sounds, including speech and music. An alternative explanation is that the auditory system is innately structured to be sensitive to small-integer frequency ratios. In any case, the auditory system presumably evolved to make use of information in the environment, including the structure of common complex sounds.

Assuming octave equivalence, each harmonic of a complex tone corresponds to a tone in the octave above the fundamental. For example, consider a tone whose fundamental is 100 Hz. The third harmonic, 300 Hz, has the same tone chroma as 150 Hz. Thus, with octave equivalence, the ratio between the fundamental and the third harmonic is 2:3 (100:150). Similarly, the fourth, fifth, sixth, seventh, and eighth harmonics stand in 1:2, 4:5, 2:3, 4:7, and 1:2 ratios to the fundamental. If the auditory system is structured to easily compare or combine very simple frequency ratios, the sixth harmonic (2:3 ratio with the fundamental)

might add more to the pitch salience of a complex tone than would the fifth harmonic (4:5 ratio with the fundamental). Similarly, the eighth harmonic might add more to the pitch salience than would the seventh (1:2 versus 4:7).

In the following experiments, the effects of harmonics on relative pitch discrimination in a melodic context are examined. In Experiment 1, the effects of sine versus complex waves and of in-tune versus out-of-tune harmonics on pitch salience are examined. In Experiment 2, the pitch salience of tones that have harmonics with simple versus more complex ratio relations to the fundamental are compared.

EXPERIMENT 1

The ability to detect a pitch change in a five-note melody based on the major triad was compared in two sets of conditions. In the first, individual tones were either sine waves or complex waves consisting of the first five harmonics. Following findings for isolated tones (Fastl & Weinberger, 1981; Platt & Racine, 1985), relative pitch discrimination was expected to be superior for sequences composed of complex over sine wave tones. In the second set, tones consisted of either an in-tune or a mistuned version of the first two harmonics. Poorer relative pitch discrimination was expected in the latter case. One study is of interest in this regard. Cohen (1982) compared melodic discrimination for sequences of tones composed of two sine waves related either by perfect fifths (2:3) or by the ratio 8:11. The latter is like a mistuned version of the former. Performance was superior with the tones related by 2:3 than 8:11.

A go/no-go vigilance task was used (e.g., see Trainor & Trehub, 1992). The standard melody repeated continuously in transposition, that is, at different pitch levels, forming a background against which changes were to be detected. Every now and then a trial occurred. On control trials, the standard melody repeated; on experimental or change trials, the third note of the melody was lowered in pitch by $\frac{1}{4}$ of a semitone. The subject's task was to respond when a change in the pitch of one note was heard on any repetition of the melody. The time at which a trial occurred was not signaled to the subject in any way. Thus, control trials were indistinguishable from the repeating background of standard melody repetitions. There were several reasons for this choice of task. First, it provides a way to maintain a strong, constant musical context. Second, although it was initially developed as an infant methodology, it has proved to be useful with adults in comparing the ease of detecting changes across different conditions (Lynch, Eilers, Oller, & Urbano, 1990; Schellenberg & Trehub, 1994a; Trainor & Trehub, 1992, 1993a, 1993b, 1994).

Method

Subjects. The subjects were 20 adults (3 male, 17 female), 19 to 36 years of age (mean = 23 years). All reported normal hearing and were free of colds at the time of testing. No adults were professional or serious amateur musicians, although some had taken music lessons (8 had 0 years, 11 had 1–5 years, and 1 had 5–10 years of lessons).

Apparatus. Participants were tested individually in an Industrial Acoustics Company sound-attenuating booth. The experiment was

controlled, and the sounds produced, by a Macintosh IIfx computer containing an Audiomeia II card (digidesign) that allowed 16-bit sound generation at a sampling rate of 22.1 kHz. The sounds were amplified by a Denon amplifier (PMA-480R) and presented through a single GSI loudspeaker (with a very flat response over frequency) located inside the booth. A response button box and feedback lights, also located inside the booth, were connected via a custom-made interface to a Strawberry Tree I/O card in the computer.

Stimuli. There were four conditions. In the sine condition, all tones were sine waves. In the complex condition, all tones consisted of the first 5 harmonics. In the in-tune condition, all tones consisted of the first two harmonics. In the out-of-tune condition, all tones consisted of Harmonic 1 and a mistuned second harmonic at $1.875 \times$ Harmonic 1. This ratio represents a semitone (i.e., $\frac{1}{12}$ th octave) mistuning. It results in a strong perception of dissonance (see, e.g., Schellenberg & Trehub, 1994b), thus increasing the chance of detecting an effect of mistuning should there be one. Within each condition, harmonics were added in sine phase at equal amplitudes.

In all other respects, the conditions were identical. A standard 5-note melody repeated in transposition (i.e., at different absolute frequencies, but with the same distances between the tones of the melody). The melody was based on the major triad, a very common form in Western musical structure (Figure 1). In the key of C major, for example, it consisted of the sequence $C_4-E_4-G_4-E_4-C_4$ (fundamental frequencies: 261.6, 329.6, 392.0, 329.6, 261.6 Hz). In the comparison melody, the third (highest) note was lowered by $\frac{1}{4}$ of a semitone (e.g., G_4 , 392.0 Hz, became 386.4 Hz). In multiharmonic conditions, all harmonics of the changed note were lowered by $\frac{1}{4}$ of a semitone as well, of course. Successive frequencies were related using equal temperament tuning. This makes transposition more straightforward, and is presumed to have little perceptual effect since even trained musicians show large variation in both interval tuning studies and performance of sequential musical intervals (Rakowski, 1990). Melody notes were contiguous and lasted 400 msec, including 10-msec rise and decay times, resulting in an onset-to-onset of 400 msec between notes. Successive melodies were separated by 1,200 msec. In order to create a good musical context, successive melodies were transposed to closely related keys (i.e., were 5 or 7 semitones apart). Starting notes of successive melodies were chosen randomly from the set B^b , F, C, G, and D (466.2, 349.2, 261.6, 392.0, 293.7 Hz), so that the above constraint was met. In other words, successive melodies started on notes that were adjacent on this list.

In the training phase (see Procedure), the standard melody repeated as in the experimental phase, but the change to be detected was much larger. For the first three training trials, the third note was raised 6 semitones (i.e., $\frac{1}{2}$ of an octave; e.g., G_4 , 391.6 Hz, became $C\#_5$, 554.4 Hz). In the remaining training trials, the third note was raised 3 semitones (i.e., $\frac{1}{4}$ of an octave; e.g., G, 391.6 Hz, became B^b , 466.2 Hz). Tones were presented at approximately 60 dB(C).

Procedure. Listeners were seated in the sound-attenuating booth opposite the experimenter, who wore headphones and listened to masking music so as to be unaware of whether a control or change trial was being presented. Throughout the test session, the standard melody repeated continuously in transposition in a quasi-random order (see stimuli), in such a way that no two consecutive repetitions were at the same pitch level. The repeating standard melody formed a constant musical context and constituted the background against which changes were to be detected. When the participant was attentive, the experimenter called for a trial by pressing a button on the response-recorder button box. Trials were not identified in any way for the experimenter or the listener. There were variable numbers of repetitions of the standard melody between trials, depending on the subject's readiness. The minimum was two, and there was no fixed maximum. On half the trials, the melody was presented with the changed note (see Stimuli); on the other half (control trials), the standard melody was presented. Thus, control trials were indistinguishable from the repeating background. Equal numbers of change and control trials were presented in quasi-random order (no more than



Figure 1. The fundamental frequencies of the standard melody in the five transpositions (change frequency indicated in parentheses).

three consecutive control trials) for a total of 24 trials per condition. Listeners were instructed to raise a hand when they detected a change. The experimenter relayed responses (hand raising) to the computer via a second button on the response box. If there was a correct response during the 1.6-sec period that began with the third (changed) note of the melody, a feedback light was illuminated. No light was illuminated for responses occurring in the absence of a change. The computer recorded all responses during the 1.6-sec response interval on both change and control trials.

The test phase of each condition was preceded by a training phase with larger changes (see Stimuli). Participants who did not meet a training criterion of four consecutive correct responses within 20 trials were eliminated. In fact, no subjects failed training in any of the experiments. All participants completed two conditions and filled out a short questionnaire about their musical background in a single test session of approximately 45 min. Ten subjects completed the sine and complex conditions and another 10 completed the in-tune and out-of-tune conditions. For each group, half the subjects received one condition first and the other half received the other condition first.

Results and Discussion

For each participant in each condition, the proportions of responses on change trials (hits) and control trials (false alarms) were transformed to d' scores according to *yes/no* tables of signal detection theory (P. V. Elliot, 1964). Occasional proportions of 0 or 1 present a problem because they result in infinite d' scores. These scores are believed to be statistically infinite rather than truly infinite (see Macmillan & Kaplan, 1985), arising from the sampling error inherent in a limited number of trials. To circumvent this difficulty, the scores were transformed prior to conversion to d' . Proportions were calculated by adding $\frac{1}{2}$ to the number of actual responses (out of 12) and dividing by the number of possible responses plus 1 (i.e., 13). This transformation maintains the original ranking of scores and has little effect on d' values (see Thorpe, Trehub, Morriongiello, & Bull, 1988). A d' of 0 represents chance performance; the maximum d' achievable under these conditions is 3.50.

As can be seen in Figure 2, top panels, relative pitch processing was superior for complex over sine wave tones, and relative pitch processing suffered when the second harmonic was mistuned. The effects of primary interest were the within-subject effects, so two separate analyses of variance (ANOVAs) were conducted with order of conditions as a between-subjects factor and either sine/complex or in-tune/out-of-tune as a within-subject factor. In both cases, there was no effect of order and no interaction involving order. There were, however, significant effects of both sine/complex [$F(1,8) = 12.63, p < .008$] and in-tune/out-of-tune [$F(1,8) = 7.78, p < .025$].

The most likely interpretation of the sine/complex effect is that the higher harmonics increase the salience of the perceived pitch. An alternative explanation is that subjects are not familiar with the sine wave timbre because sine waves occur rarely in the natural environment, and this affected their ability to compare tones with this timbre. However, it might be argued that subjects are also not familiar with the 5-harmonic complex where the components are added in sine phase at equal amplitudes. In fact, the sine wave stimuli sound subjectively close to a flute timbre, so degree of familiarity with the timbre seems an unlikely explanation.

As can be seen in Figure 2, top panels, and confirmed in a t test, there was no significant difference between performance on the complex condition, where tones consisted of the first 5 harmonics, and performance on the in-tune condition, where tones consisted of the first 2 harmonics. While it might be tempting to conclude that Harmonics 3, 4, and 5 did not contribute to pitch salience, such a conclusion is probably unwarranted because the power of this test is low. Different subjects were tested in each condition, and there were relatively few subjects (10) in each condition who completed relatively few trials (24). (The smaller the number of trials, the greater the error variance.) The top panels in Figure 2 suggest that performance on the out-of-tune condition might have been better than performance on the sine condition, but this difference is also not signifi-

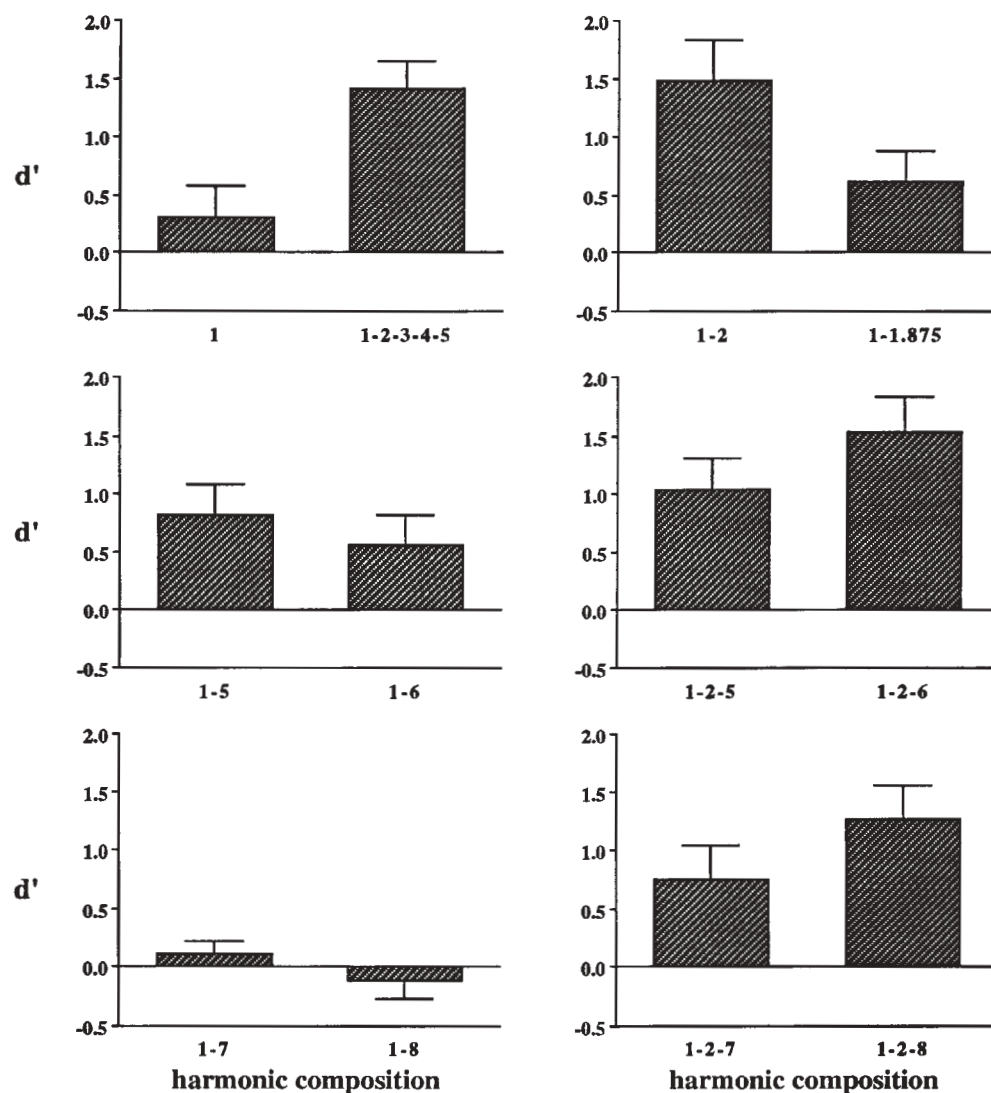


Figure 2. d' performance by condition: top left panel, sine versus complex (Experiment 1). Top right panel, in-tune versus out-of-tune (Experiment 1). Middle panels, 1-5 versus 1-6 and 1-2-5 versus 1-2-6 (Experiment 2). Bottom panels, 1-7 versus 1-8 and 1-2-7 versus 1-2-8 (Experiment 2). Error bars represent the standard error of the mean.

cant. However, t tests revealed that performance on out-of-tune was above chance levels [$t(9) = 2.3, p < .03$] (although if the number of t tests performed is taken into consideration, this effect is also not significant at .05). Performance on sine was not significantly above chance levels.

Musical training (i.e., number of years of music lessons) was not significantly correlated with performance on any condition, or with the difference in performance between the sine/complex or in-tune/out-of-tune conditions. However, no subjects were professional musicians, and the number of years of music lessons is not necessarily a good index of musical ability, so this null finding should not be taken to confirm that musical experience or ability has no effect on relative pitch processing with these stimuli.

EXPERIMENT 2

In this experiment, the question of whether harmonics standing in octave-equivalent octave and perfect fifth relationships to the fundamental would facilitate pitch perception more than harmonics standing in more complex relationships was examined. The eighth harmonic is three octaves above the fundamental. The seventh harmonic, when reduced to its octave-equivalent frequency in the octave above the fundamental, stands in a ratio of 4:7 with the fundamental. Thus, it was expected that relative pitch discrimination would be superior when tones contained Harmonic 8 over Harmonic 7. The sixth harmonic, when reduced to its octave-equivalent frequency in the octave

above the fundamental, stands in a 2:3 ratio with the fundamental, and the fifth harmonic in a 4:5 ratio. Similarly, performance was expected to be superior when tones contained Harmonic 6 over Harmonic 5.

In order to test the differential effects of these harmonics, it was desirable to use widely spaced harmonics, as there is evidence that adjacent harmonics can interact (Moore et al., 1985; Moore, Glasberg, & Shailer, 1984). Thus, tones consisting of the fundamental and one of Harmonics 5, 6, 7, or 8 were used. Most experimental studies, as well as theories (e.g., Gerson & Goldstein, 1978; Goldstein, 1973; Terhardt, 1974; Terhardt et al., 1982; Wightman, 1973), of pitch perception have focused on complex sounds with successive harmonics. While the pitch of complexes of non-successive harmonics can be ambiguous in the absence of the fundamental (see, e.g., Schouten, Ritsma, & Cardozo, 1962), subjects are able to track the missing fundamental in a musical interval task (Houtsma, 1979; Gerson & Goldstein, 1978). In the present experiments, the fundamental was always present.

However, tones consisting of such widely spaced harmonics tend to segregate perceptually into their sine wave components (Bregman, 1990; Van Noorden, 1975). Thus, in considering how different harmonics interact with the fundamental in providing input to pitch processors, it is also informative to manipulate how readily the components of each tone fuse into a single percept. In general, the melodic context, in which the timbre of successive tones remains the same (i.e., all components move in parallel between successive notes) would be expected to promote fusion (Bregman, 1990; McAdams, 1984). However, the addition of more harmonics would also be expected to promote fusion. It was hypothesized that the addition of the second harmonic would increase the integration of each of these stimuli, and this hypothesis was tested in a mini-experiment.¹ In a two-interval forced choice design, each trial consisted of two versions of the major triad melody from Experiment 1, where the versions differed in the timbre of the tones. Subjects, who had had some experience performing auditory psychoacoustic tasks, were told that all tones contained a high, isolated harmonic. They were asked to indicate whether it was easier to hear out the high harmonic as a separate tone in the first or second melody. Three subjects completed all four conditions. Each condition consisted of 24 trials. Two conditions were at the highest fundamental frequency level to occur in the following experiment ($B\flat_4$ - D_5 - F_5 - D_5 - $B\flat_4$; 466.2, 554.4, 659.3, 554.4, 466.2 Hz) and two were at the lowest (C_4 - E_4 - G_4 - E_4 - C_4 ; 261.6, 329.6, 392.0, 329.6, 261.6 Hz). Crossed with this factor, in one condition, melodies consisting of tones containing Harmonics 1 and 5 were compared to melodies consisting of tones containing Harmonics 1, 2, and 5. In the other condition, melodies consisting of tones containing Harmonics 1 and 8 were compared to tones containing 1, 2, and 8. Each subject in each condition chose the melody whose tones did not contain the second harmonic as easiest to hear out the high harmonic between 92% and 100% of the time. Thus, the presence of the second har-

monic led to significantly greater fusion of the harmonics into a single percept.

In Experiment 2, therefore, the contribution of various harmonics to relative pitch discrimination under conditions of more or less fusion was examined by comparing conditions where the tones consisted of the following harmonics: 1 and 8 versus 1 and 7; 1, 2, and 8 versus 1, 2, and 7; 1 and 6 versus 1 and 5; 1, 2, and 6 versus 1, 2, and 5.

Method

Subjects. The subjects were 40 adults (13 male, 27 female), 19 to 30 years of age (mean = 22 years). No adults were professional or serious amateur musicians, although some had taken music lessons (12 had 0 years, 15 had 1–5 years, and 13 had 5–10 years of lessons).

Stimuli. The melodies and changes were identical to those of Experiment 1. However, the harmonic structure of the tones was different. The fundamental was present in all conditions. Two factors varied across conditions: the simplicity of the ratio between the high harmonic and the fundamental, and whether or not the second harmonic was present (affecting how fused the components were perceived to be). Conditions will be identified according to the harmonic structure of the tones: 1-5; 1-6; 1-7; 1-8; 1-2-5; 1-2-6; 1-2-7; 1-2-8.

Apparatus. The apparatus was identical to that of Experiment 1.

Procedure. Ten subjects completed both 1-5 and 1-6; 10 completed both 1-7 and 1-8; 10 completed both 1-2-5 and 1-2-6; and 10 completed both 1-2-7 and 1-2-8. In each case, half the subjects received one condition first and half received the other condition first. Within each condition, the procedure was identical to that of Experiment 1.

Results and Discussion

Performance was generally better when Harmonics 5 or 6 were present than when Harmonics 7 or 8 were present and when Harmonic 2 was present than when it was absent (Figure 2, two lower rows). The predicted effects of ratio complexity occurred only when Harmonic 2 was present (i.e., when the tones fused into a single percept). Under this condition, relative pitch discrimination was superior when tones contained harmonics standing in octave and perfect fifth relations to the fundamental. Hit and false alarm rates were transformed to d' scores as in Experiment 1. An overall repeated measures ANOVA was conducted with low/high harmonics (i.e., 1-5, 1-6, 1-2-5, 1-2-6 vs. 1-7, 1-8, 1-2-7, 1-2-8), degree of fusion (i.e., 1-5, 1-6, 1-7, 1-8 vs. 1-2-5, 1-2-6, 1-2-7, 1-2-8), and order of conditions (i.e., 1-5, 1-7, 1-2-5, or 1-2-7 first vs. 1-6, 1-8, 1-2-6, or 1-2-8 first) as between-subjects factors and ratio complexity (i.e., 1-5, 1-7, 1-2-5, 1-2-7 vs. 1-6, 1-8, 1-2-6, 1-2-8) as a within-subject factor. There was a significant effect of low/high harmonics [$F(1,32) = 4.27, p < .05$]. In general, the lower the harmonic, the greater its influence on the pitch. Degree of fusion was also significant [$F(1,32) = 11.95, p < .002$], with higher performance when Harmonic 2 was present than when it was absent. There was no effect of order or any interactions with order. The main effect of ratio complexity was not significant, but the interaction between ratio complexity and degree of fusion was significant [$F(1,32) = 11.68, p < .002$]. This interaction was explored further by conducting separate analyses for the conditions in which Harmonic 2 was present and those in which it was absent.

For the less fused stimuli, an ANOVA with low/high harmonics (i.e., 1-5, 1-6 vs. 1-7, 1-8) as a between-subjects factor and ratio complexity (i.e., 1-5, 1-7 vs. 1-6, 1-8) as a within-subject factor revealed a significant effect only of low/high harmonics [$F(1,18) = 7.41, p < .02$]: Performance was superior with the lower harmonics. There was no effect of ratio complexity. On the other hand, for the more fused stimuli, an ANOVA with low/high harmonics (i.e., 1-2-5, 1-2-6 vs. 1-2-7, 1-2-8) as a between-subjects factor and ratio complexity (i.e., 1-2-5, 1-2-7 vs. 1-2-6, 1-2-8) as a within-subject factor revealed only an effect of ratio complexity [$F(1,18) = 11.71, p < .003$], with superior performance when the ratio of the harmonic to the fundamental was simpler (Figure 2, two lower rows).

Comparing the results of Experiments 1 and 2 (Figure 2) suggests that adding higher harmonics either does not affect performance or might actually depress it. For example, performance did not differ significantly across Conditions 1-2 (in-tune, Experiment 1), 1-2-6 (Experiment 2) and 1-2-8 (Experiment 2). However, it appears that performance was worse on 1-2-5 (Experiment 2) and 1-2-7 (Experiment 2) than it was on 1-2 (in-tune, Experiment 1), although these differences were in fact not significant by *t* tests. Performance on 1-7 (Experiment 2) and 1-8 (Experiment 2) appeared to be lower than performance on 1 (sine, Experiment 1), but performance on 1-5 (Experiment 2) and 1-6 (Experiment 2) appeared to be higher than performance on 1 (sine, Experiment 1). However, none of these effects were statistically significant. It should be kept in mind that the power to discriminate these potential differences is relatively low: These are between-subjects effects, with only 10 subjects per condition who completed only 24 trials each. The experiment was designed so the effects of primary interest were within-subject. Of these, performance on 1-2-6 was superior to performance on 1-2-5, and performance on 1-2-8 was superior to performance on 1-2-7. On the basis of these data, then, it is not possible to identify which of the following two interpretations is more likely: When the harmonics fuse into a single percept, (1) adding Harmonic 6 or 8 increases pitch salience more than adding Harmonic 5 or 7, or (2) adding Harmonic 6 or 8 has no effect on pitch salience whereas adding Harmonic 5 or 7 decreases pitch salience.

Performance on 1-7 and 1-8 did not differ significantly from performance on 1 (sine), and performances on 1-7 and 1-8 were not above chance levels, again suggesting that these conditions were particularly difficult. Further, performance on 1-1.875 (out-of-tune) was superior to performance on 1-8 [$t(18) = 2.37, p < .03$] (although, again, if the number of *t* tests is taken into consideration, this effect is not significant at .05). Thus the presence of this mistuned second harmonic appeared to add more to pitch salience than did the presence of the eighth harmonic when it did not fuse with the fundamental.

Again there were no significant correlations between the number of years of music lessons and performance on any condition, or the number of years of music lessons and the difference between performance on any two conditions completed by any subject. Further, an ANOVA comparing

the number of years of music lessons across the six groups of subjects in Experiments 1 and 2 revealed no significant differences in musical training.

GENERAL DISCUSSION

Experiment 1 showed that pitch discrimination in a musical context was superior when tones were composed of the first five harmonics than when tones were sine waves. Further, performance was higher when tones were composed of an in-tune version of the first two harmonics than when tones were composed of an out-of-tune version. Thus these findings generalize to pitch perception in a musical context: Pitch salience increases with the addition of harmonics and decreases when harmonics are out of tune.

Experiment 2 showed that under conditions where the harmonics fused into a single percept, pitch discrimination was superior when Harmonic 8 was present over Harmonic 7, and when Harmonic 6 was present over Harmonic 5. Thus, pitch salience did not decrease strictly with increasing harmonic number. Under these conditions, the presence of harmonics standing in octave or perfect fifth relations to the fundamental appear to have resulted in greater pitch salience than when other nearby harmonics were present. When the fundamental and harmonics fused less into a single percept, there was no effect of the simplicity of frequency ratio between the fundamental and harmonics. Rather, pitch salience simply decreased with increasing harmonic number.

It seems unlikely that combination tones arising from distortions in the inner ear could be responsible for these effects. The most common combination tones of two frequencies, f_{high} and f_{low} , are of the form $f_{\text{high}} - f_{\text{low}}$ or $nf_{\text{low}} - (n-1)f_{\text{high}}$ (Plomp, 1965; Roederer, 1979). The formula $nf_{\text{low}} - (n-1)f_{\text{high}}$, when used with widely spaced harmonics, results in negative numbers. Combination tones of the form $f_{\text{high}} - f_{\text{low}}$ occurring between the high harmonic of interest and the fundamental would simply be one harmonic lower than the high harmonic. In fact, all combination tones arising from the tones used in these experiments would occur on harmonics or subharmonics of the fundamental. Further, widely spaced tones give rise to fewer, less intense combination tones and, for the most part, only combination tones below the lowest tone are heard (Plomp, 1965). It seems likely, then, that the enhanced contributions of the sixth and eighth harmonics to pitch perception arise more centrally in the auditory system.

The stimuli employed in Experiment 2 were peculiar in that the harmonics were widely spaced. The use of widely spaced harmonics minimizes the influence of neighboring harmonics on the harmonic of interest, but such stimuli are rare indeed in the natural environment. The wide spacing may be related to some of the surprising results of comparisons across Experiments 1 and 2. For example, the addition of either Harmonic 7 or 8 to a sine wave, or the addition of Harmonic 5 or 6 to a fundamental plus second harmonic appears either to have had no effect or to actually have decreased pitch salience. Further, performance was superior with tones composed of a fundamental plus second har-

monic mistuned by a semitone than with a fundamental and eighth harmonic. Thus widely spaced harmonics appear to lead to pitch processing difficulty in general.

The major models of pitch perception (e.g., the optimum processor theory, Gerson & Goldstein, 1978; Goldstein, 1973; the virtual pitch theory, Terhardt, 1974; Terhardt et al., 1982) are based largely on results of highly controlled experiments with well-fused tones and impoverished contexts. Thus, until it is known whether the ratio simplicity results generalize to a variety of contexts, and until the relation between fusion and pitch salience is determined, it is premature to examine models of pitch perception in great detail with respect to the present results. However, Houtsma (1979) pointed out that the optimum processor model (Gerson & Goldstein, 1978) did not capture the domination of the octave harmonics on perceived pitch with widely spaced harmonics. Should the results of the present investigation generalize to different contexts, the prior probabilities of the optimum processor model could be modified to incorporate these effects.

The virtual pitch model (Terhardt, 1974, 1984; Terhardt et al., 1982) was tested with stimuli from the present experiments. The virtual pitch model performs an initial spectral analysis, yielding a set of *spectral pitches*, including effects of combination tones and pitch shifts in components due to level and masking effects. *Virtual pitches* are integer factors of the spectral pitches. They are weighted according to the number of components contributing to that virtual pitch, the weight of the spectral pitches from which they arise, the subharmonic number, and the accuracy of nearly virtual pitch coincidences. The virtual pitches of the various timbres used in the experiments reported here were calculated using the VPITCHu3 computer program.² For a given set of harmonics (i.e., frequency and dB listings), the program produces a set of spectral and virtual pitches, each with a frequency and weight. The frequency associated with the largest weight represents the pitch most likely to be perceived, and the weights can be considered measures of the pitch salience. A mid-frequency fundamental (466 Hz) was used. Sine wave tones were at 60 dB and each component of two-component stimuli was at 57 dB; each component of three-component stimuli was at 55.3 dB; and each component of five-component stimuli was at 53 dB.

The virtual pitch weightings produced by the model were in general agreement with the present experimental results in that the more harmonics present, the more salient the pitch, and the presence of lower harmonics resulted in higher pitch salience than the presence of higher harmonics. In addition, the superiority of the in-tune over the out-of-tune combination was reflected in the obtained ratings. However, the model predicted a large jump in pitch salience from Complex 1-2 to Complex 1-2-3-4-5 (virtual pitch weightings of .60 and 1.70, respectively), but performance with these timbres was equivalent in the present experiments. Interestingly, for Complexes 1-5, 1-6, 1-7, and 1-8, spectral pitches at the two component frequencies had much higher weights than did any virtual pitches (e.g., for 1-5, spectral pitch weights were .47 at 466 Hz, and .38 at

2330 Hz, whereas the highest virtual pitch weight was .19 at 233 Hz). This is in agreement with the present finding that these complexes did not fuse readily into one percept with one pitch. The model did not capture the ratio simplicity effects. Virtual pitch weights decreased from Complex 1-2-5 to Complex 1-2-6 to Complex 1-2-7 to Complex 1-2-8 (weights were .94, .89, .86, and .83, respectively), whereas performance in Experiment 2 decreased from Condition 1-2-6 to Condition 1-2-8 to Condition 1-2-5 to Condition 1-2-7. If further studies reveal that the harmonic ratio simplicity effects generalize to a variety of contexts, the weights in Terhardt's model could be adjusted to yield appropriate results.

An alternative modeling approach would be to start with an architecture that incorporates small-integer ratios into its basic design, such as Patterson's (1986) spiral pulse-stream processor, whereby equally spaced pulses (nerve firings) flow along a time line that is wrapped into a spiral whose length doubles on each successive circuit.

A number of questions arise from the present results. First, do the effects of harmonic ratio simplicity generalize to impoverished or nontonal contexts? Second, do the effects of harmonic ratio simplicity also generalize to more natural timbres (i.e., without widely spaced harmonics)? Finally, following suggestions that the perceived consonance of simultaneous tones is related to their degree of fusion (see DeWitt & Crowder, 1987), are complexes with harmonics standing in simple ratios to the fundamental (e.g., 1-2-6 and 1-2-8 of Experiment 2) also perceived as more fused (as well as having a better defined pitch) than complexes with harmonics in more complex ratios (e.g., 1-2-5 and 1-2-7 of Experiment 2)?

REFERENCES

- AYRES, T., AESCHBACH, S., & WALKER, E. L. (1980). Psychoacoustic and experiential determinants of tonal consonance. *Journal of Auditory Research*, **20**, 31-42.
- BREGMAN, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
- BURNS, E. M., & WARD, W. D. (1982). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (pp. 241-269). New York: Academic Press.
- COHEN, A. J. (1982). Exploring the sensitivity to structure in music. *Canadian University Music Review*, **3**, 15-30.
- COHEN, A. J., THORPE, L. A., & TREHUB, S. E. (1987). Infants' perception of musical relations in short transposed tone sequences. *Canadian Journal of Psychology*, **41**, 33-47.
- DEMANY, L., & ARMAND, F. (1984). The perceptual reality of tone chroma in early infancy. *Journal of the Acoustical Society of America*, **76**, 57-66.
- DEWITT, L. A., & CROWDER, R. G. (1987). Tonal fusion of consonant musical intervals: The oomph in Stumpf. *Perception & Psychophysics*, **41**, 73-84.
- DIVENYI, P. L., & HIRSH, I. J. (1974). Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America*, **56**, 144-151.
- DOWLING, W. J., & HARWOOD, D. L. (1986). *Music cognition*. Orlando, FL: Academic Press.
- ELLIOT, J., PLATT, J. R., & RACINE, R. J. (1987). Adjustment of successive and simultaneous intervals by musically experienced and inexperienced subjects. *Perception & Psychophysics*, **42**, 594-598.
- ELLIOT, P. V. (1964). Tables of d' . In J. A. Swets (Ed.), *Signal detection*

- and recognition by human observers: *Contemporary readings* (Appendix 1, pp. 651-684). New York: Wiley.
- FASTL, H., & WEINBERGER, M. (1981). Frequency discrimination for pure and complex tones. *Acoustica*, **20**, 521-534.
- GERSON, A., & GOLDSTEIN, J. L. (1978). Evidence for a general template in central optimal processing for pitch of complex tones. *Journal of the Acoustical Society of America*, **63**, 498-510.
- GOLDSTEIN, J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, **54**, 1496-1516.
- GOULD, J. L., & MARLER, P. (1987). Learning by instinct. *Scientific American*, **256**, 74-85.
- GREENWOOD, D. D. (1991). Critical bandwidth and consonance in relation to cochlear frequency-position coordinates. *Hearing Research*, **54**, 164-208.
- HOUTSMA, A. J. M. (1979). Musical pitch of two-tone complexes and predictions by modern pitch theories. *Journal of the Acoustical Society of America*, **66**, 87-99.
- KAMEOKA, A., & KURIYAGAWA, M. (1969a). Consonance theory part 1: Consonance of dyads. *Journal of the Acoustical Society of America*, **45**, 1451-1459.
- KAMEOKA, A., & KURIYAGAWA, M. (1969b). Consonance theory part 2: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America*, **45**, 1460-1469.
- KOLINSKI, M. (1967). Recent trends in ethnomusicology. *Ethnomusicology*, **11**, 1-24.
- LEVELT, W. J. M., VAN DE GEER, J. P., & PLOMP, R. (1966). Triadic comparisons of musical intervals. *British Journal of Mathematical & Statistical Psychology*, **19**, 163-179.
- LYNCH, M. P., EILERS, R. E., OLLER, D. K., & URBANO, R. C. (1990). Innateness, experience, and music perception. *Psychological Science*, **1**, 272-276.
- MACMILLAN, N. A., & KAPLAN, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, **98**, 185-199.
- MCADAMS, S. (1984). *Spectral fusion, spectral parsing, and the formation of auditory images*. Unpublished doctoral dissertation, Stanford University.
- MOORE, B. C. J., GLASBERG, B. R., & PETERS, R. W. (1985). Relative dominance of individual partials in determining the pitch of complex tones. *Journal of the Acoustical Society of America*, **77**, 1853-1860.
- MOORE, B. C. J., GLASBERG, B. R., & SHAILER, M. J. (1984). Frequency and intensity difference limens for harmonics within complex tones. *Journal of the Acoustical Society of America*, **75**, 550-561.
- NETTL, B. (1956). *Music in primitive culture*. Cambridge, MA: Harvard University Press.
- PATTERSON, R. D. (1986). Spiral detection of periodicity and the spiral form of musical scales. *Psychology of Music*, **14**, 44-61.
- PLATT, J. R., & RACINE, R. J. (1985). Effect of frequency, timbre, experience, and feedback on musical tuning skills. *Perception & Psychophysics*, **38**, 543-553.
- PLOMP, R. (1965). Detectability thresholds for combination tones. *Journal of the Acoustical Society of America*, **37**, 1110-1123.
- PLOMP, R. (1967). Pitch of complex tones. *Journal of the Acoustical Society of America*, **41**, 1526-1533.
- PLOMP, R., & LEVELT, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, **38**, 548-560.
- RAKOWSKI, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music*, **18**, 60-72.
- RITSMA, J. R. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *Journal of the Acoustical Society of America*, **42**, 191-198.
- ROEDERER, J. G. (1979). *Introduction to the physics and psychophysics of music* (2nd ed.). New York: Springer-Verlag.
- SHELLENBERG, E. G., & TREHUB, S. E. (1994a). Frequency ratios and the discrimination of pure tone sequences. *Perception & Psychophysics*, **56**, 472-478.
- SHELLENBERG, E. G., & TREHUB, S. E. (1994b). Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin & Review*, **1**, 191-201.
- SCHOUTEN, J. F., RITSMA, R. J., & CARDOZO, B. L. (1962). Pitch of the residue. *Journal of the Acoustical Society of America*, **34**, 1418-1424.
- TERHARDT, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, **55**, 1061-1069.
- TERHARDT, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music Perception*, **1**, 276-295.
- TERHARDT, E., STOLL, G., & SEEWANN, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, **71**, 679-688.
- THORPE, L. A., TREHUB, S. E., MORRONGIELLO, B. A., & BULL, D. (1988). Perceptual grouping by infants and preschool children. *Developmental Psychology*, **24**, 484-491.
- TRAINOR, L. J. (1991). *The origins of musical pattern perception: A comparison of infants' and adults' processing of melody*. Unpublished doctoral dissertation, University of Toronto.
- TRAINOR, L. J., & TREHUB, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 394-402.
- TRAINOR, L. J., & TREHUB, S. E. (1993a). Musical context effects in infants and adults: Key distance. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 615-626.
- TRAINOR, L. J., & TREHUB, S. E. (1993b). What mediates adults' and infants' superior processing of the major triad? *Music Perception*, **11**, 185-196.
- TRAINOR, L. J., & TREHUB, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception & Psychophysics*, **56**, 125-132.
- VAN NOORDEN, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Technische Hogeschool Eindhoven, Eindhoven, The Netherlands.
- VOS, J. (1986). Purity ratings of tempered fifths and major thirds. *Music Perception*, **3**, 221-258.
- WIGHTMAN, F. L. (1973). The pattern-transformation model of pitch. *Journal of the Acoustical Society of America*, **54**, 407-416.

NOTES

1. This mini-experiment was conducted at the suggestion of the reviewers of an earlier draft of this manuscript.
2. Copyright: E. Terhardt, Institute of Electroacoustics, Technical University, D-8000 Munich 2, Germany.

(Manuscript received May 31, 1994;
revision accepted for publication October 15, 1995.)