The influence of spectral composition of complex tones and of musical experience on the perceptibility of virtual pitch

ANNEMARIE PREISLER
Karl-Franzens-University, Graz, Austria

A matching paradigm was used to evaluate the influence of the spectral characteristics number, relative height, and density of harmonics on the perceptibility of the missing fundamental. Fifty-eight musicians and 58 nonmusicians were instructed to adjust mistuned sinusoids to the subjectively perceived fundamental pitches of corresponding overtone spectra. Analyses of variance were used to compare the average of absolute and relative deviations of the tunings from the highest common divisors of the complex tones. The results indicate that musical experience is the most influential single factor determining the assessment of fundamental pitch. Nevertheless, all spectral parameters significantly affect tuning performance. Systematic relative deviations (stretching/compression effects) were observed for all considered variables. An increase of the optimum subjective distance between an overtone spectrum and its corresponding fundamental was characteristic of musicians and unambiguous spectra, whereas the compression effect was typical of nonmusicians and complex tones containing spectral gaps.

Since Seebeck (1843) discovered that humans are able to perceive the pitch of a complex tone even though the tone does not contain any spectral component corresponding to that pitch, the perception of the *missing fundamental* has been the subject of various experimental and theoretical approaches. Over time, it has been observed that this phenomenon is simply a special case of virtual pitch perception, whereby the fundamental pitch assigned to a complex tone is more strongly determined by higher harmonics than by the fundamental frequency itself.

Fletcher (1929, 1931) described the effect as an auditory illusion caused by combination tones originating from the middle or the inner ear. (Combination tones result from nonlinear distortion and can be measured objectively within the cochlea.) However, various observations proved that this was not true. For example, no beats are perceived if a real tone with a frequency slightly different from that of the pitch is added to the stimulus (Schouten, 1940b). Furthermore, the pitch cannot be masked by bandpass noise in the region of the missing fundamental (Patterson, 1969).

Schouten (1938, 1940a, 1940b) assumed that "residue pitch" is caused by high harmonics that cannot be resolved by the inner ear and are thus superimposed on the basilar membrane within the range of bandpass filters. Theoret-

ically, a pitch evoked by the periodicity of the envelope of the rest of the unresolved partials should correspond to that of the missing fundamental. This model, based on the restricted filter characteristics of the inner ear, was shown to be insufficient by the discovery of the *principle of dominance* first described by Ritsma (1967) and Plomp (1967). These authors showed that the lower partials, which are resolved well in the cochlea and whose pitch is determined by the maximum locus of excitation on the basilar membrane, have a stronger effect on the perceptibility of virtual pitch than do the higher partials. This finding led to the assumption that there is an underlying central processor, relating a fundamental pitch to the frequencies of aurally resolved partials according to criteria of the highest probability (Goldstein, 1973; Terhardt, 1974, 1976). In contrast to the periodicity model, this hypothetical pitch processor is supposed to be operating at a higher neuronal level and thus to be sensitive to learning experience.

De Boer's (1956a, 1956b) experiments on fundamental pitch perception of inharmonic sound spectra provided new evidence for the importance of resolved partials. When the part tones of a complex tone are shifted by the same amount, the periodicity of the envelope remains the same while the fine structure of the stimulus is changed. Under this condition, it is the distance between the two maxima of the carrier frequency—that is, the fine structure—rather than the distance between the two maxima of the modulation frequency that corresponds to the perceived fundamental pitch. Consequently, De Boer proposed a "pitch picking mechanism" that would extract pitch from successive maxima of the carrier frequency within the envelope. Wightman (1973) pointed out that such a mechanism

This research was supported by the Austrian Funds for Scientific Research (FWF) as Project SPR 8060. The entire study was carried out in the Department of Musicology, Karl-Franzens-University, Graz, Austria. The author can now be reached at the Department of Zoology, Ludwig-Maximilians-University, Luisenstrasse 14, 80333 Munich 2, Germany.

would have to be extremely phase sensitive. In his experiments, he showed that phase changes have no effect on the perceived pitch. Houtsma and Goldstein (1971, 1972) found that the missing fundamental was still perceived by musicians when only two harmonic components were presented to different ears. Although the perceived binaural pitch was reported to be rather weak, the sensation cannot be explained by models that attribute the virtual pitch phenomenon to peripheral mechanisms only.

Generally, time-coding models refer to the principle of neural volleying (Wever, 1949) and to the fact that the temporal structure of a tone is preserved up to 5 kHz within the eighth nerve in mammals (Anderson, Rose, Hind, & Brugge 1971; Javel, 1986; Kiang, Watanabe, Thomas, & Clark, 1965; Pfeiffer & Molnar, 1970). As the pattern of action potentials is correlated to the phase of a periodic sound wave, the effect is referred to as *phase locking*. The ability of neurones to respond according to the phase of the stimulus decreases rapidly within the ascending neural pathway.

Langner and Schreiner's (1988) results should be mentioned as representing an important attempt to establish a neurophysiological substrate for a time-coding principle. The authors report the presence of a tonotopic arrangement of cells within the central nucleus of the inferior colliculus of the cat for modulation frequencies between 100 and 500 Hz.

Recently, Sedlmeier (1992) showed that bats can be trained to react as if they perceived a missing fundamental within the ultrasonic region (20–25 kHz). Unfortunately, the possibility of subjects' having learned the behavior from spectral clues in these experiments cannot be excluded. If the virtual aspect of information actually gave rise to the observed behavior, this finding cannot be explained by time-coding models, since phase locking is restricted to much lower frequencies (Javel, 1986).

Supporters of time concepts frequently cite the pitch of interrupted white noise as a crucial effect (Miller & Taylor, 1948). Such spectra evoke a weak pitch sensation, although they do not contain peaks at a certain frequency. Hence it can be concluded that the perceived pitch is the result of the modulation frequency only. Gruber (1992) pointed out that the obtained results might be an artefact caused by slight deviations of the spectrum from pure white noise.

A new version of the time concept was proposed by Meddis and Hewitt (1991). Referring to an earlier concept elaborated by Licklider (1951), these authors developed a computer simulation model on the basis of autocorrelation functions. Its most striking difference from previous concepts is the idea that the information, derived by different cochlear filter channels, is combined across channels at a higher level of neuronal processing. Thus, the model is able to account for binaural effects, for the pitch shift of inharmonic signals, and for the predominance of lower partials. The "principle of dominance," a former point of criticism directed at time models, is explained by the higher density of filter channels across the basilar

membrane within the lower frequency region. It is claimed that the involvement of more filters accounts for the fact that low partials are weighted to a greater extent than high partials are. At present, time and place models are able to explain the bulk of psychoacoustical data. Evidence from recent cochlear implant research on deaf subjects indicates that both principles might be used for pitch coding (e.g., Eddington, 1980).

Physiological Versus Psychological Theories of Pitch Perception

The validity of spectral/temporal models has interesting implications for the question of the extent to which harmonic perception is determined by physiological or by environmental factors. A mechanism using neuronal discharge patterns for the analysis of a harmonic complex tone should always produce a distinct fundamental pitch—at least for signals whose components are all in cosine phase (Ritsma & Engel, 1964). Such a criterion for the most probable solution does not necessarily exist for pure place concepts. In principle, models taking advantage of place coding allow that any frequently occurring stimulus configuration can be learned with the same probability.

Terhardt (1972, 1978, 1982) proposed that the fundamental pitch extraction is learned in early infancy through the memorization of patterns of voiced sounds of speech that contain harmonic components (Stoll, 1980). At present, however, the results from tests on infant virtual pitch perception are controversial. Bundy and Colombo (1982) found no indication of the perception of the missing fundamental in 4-month-olds; Clarkson and Clifton (1984) reported that infants at the age of 7–8 months already displayed virtual pitch categories. This discrepancy might be due to differences in experimental method, or it may indicate that the ability of fundamental pitch perception is acquired between the 4th and 7th months.

Some investigations have addressed the question of generalizability of the virtual pitch phenomenon across different species of vertebrates. Heffner and Whitfield (1976) reported that cats are able to perceive the *missing fundamental*. Similar results were published for guinea fowl (Langner, 1983), rhesus monkeys (Tomlinson & Schwarz, 1988), starlings (Cynx & Shapiro, 1986), and bats (Sedlmeier, 1992).

Generally, such results demonstrate that fundamental pitch perception is common among nonhuman species and humans of different ages, so that it seems to have evolved under the pressure of natural selection. Nevertheless, these comparative results do not address the questions of whether and to what extent this pressure has been applied to physiological mechanisms or learning abilities. Theoretically, the ability in question may be determined by genetic factors; acquired in early childhood through learning predispositions; or even acquired throughout life, from extensive experience with harmonic stimulus material.

From Terhardt's (1972, 1978, 1982) point of view, the pressure of natural selection has affected speech acquisition abilities, which, as a consequence, have generated

the by-product of "musical abilities." If "speech" is understood in a more general sense to include acoustical communication among animals, this concept may be applied to different species whose voice production mechanisms are similar to those of humans. In more recent papers, Terhardt (1989, 1991) has described fundamental pitch perception as being closely related to processes of acquisition of information within the visual modality of perception. Spectral pitches are referred to as *acoustical contours* that shape the final gestalt percept of virtual pitch. In analogy to partly covered visual shapes that are spontaneously completed in the mind of the observer, virtual pitch perception is understood to be a basic feature of acoustical perception, providing important clues for signal-source identification and separation.

A question that hitherto has not been investigated in experimental research is to what extent the ability to perceive the missing fundamental can be developed throughout life through musical education. Because it is problematic to draw causal conclusions from correlational data only, additional longitudinal studies are needed, to provide clues for the interpretation of interindividual differences. Nevertheless, a great variability among subjects, regardless of its actual source, can be explained by a more flexible underlying mechanism. This seems legitimate, for it is commonly accepted in psychobiology that features varying to a greater extent among the population are determined by essential physiological functions only to a small degree. If fundamental pitch perception is first and foremost due to a rigid subcortical temporal autocorrelation mechanism, this basic feature should exhibit less variation among different subjects than would a mechanism sensitive to learning experience. In the present study, the differences between musicians and nonmusicians were analyzed according to this philosophy.

The spectral characteristics of the complex tone stimuli were modified to compare the degree of variability evoked by physical parameters with the degree of variability induced by characteristics of the perceiving subjects. Interactions between stimulus-related and subject-related variables are discussed.

METHOD

Apparatus and Procedure

The experimental apparatus comprised a computer (Atari Mega STE) and a sound sampler (Akai S1100).

The stimuli were presented monaurally on closed headphones (AKG K270). The stimulus material for the virtual pitch demonstration and the tuning test was generated by a commercial additive synthesis program (Turbosynth, by Digidesign) and stored at the sampler. Instructions were presented on an Atari monochrome monitor located at eye level in front of the subject. The arrangement is shown in Figure 1.

Subjects

Fifty-eight musicians and 58 nonmusicians participated in the experiments. Their degree of musical training was assessed by questionnaire. The subjects had to indicate whether or not they had played one or more instruments, how long they had done so, and how much

Figure 1. Experimental arrangement.

time they usually spent in practicing. For each instrument, the period of playing was multiplied by the average time of practicing per day, week, or month, and the scores were finally summed up.

All individuals who indicated that they had played one or more instruments for less than 3 years were considered as nonmusicians.

In addition, the subjects had to evaluate their own musical abilities by indicating their characteristic musical behavior and history. They were asked whether they considered themselves to be (1) professional musicians, (2) hobby musicians, or (3) occasional players; whether they had learned to play their instrument(s) (1) through professional musical education, (2) through semiprofessional education, or (3) autodidactically; and whether they were able to read music at sight (1) easily, (2) hardly, or (3) not at all. Category 1 answers were assigned 3 points; Category 2 answers, 2 points; and Category 3 answers, 1 point.

All subjects who indicated that they had played an instrument regularly for at least 3 years (i.e., at least 1,095 h per year, corresponding to an average duration of 3 h per day), were included in the musicians category. Subjects who indicated they had played for at least 3 years frequently (i.e., at least 365 h per year, corresponding to 1 h per day) were also assigned to the group of musicians, provided that they had scored above 6 on the three self-evaluation musical history questions. The remaining participants were classified as nonmusicians.

All subjects were tested audiometrically before the experimental sessions began. Subjects with a hearing loss of at least 20 dB SPL at any test frequency up to 8 kHz were excluded from further experiments.

Virtual Pitch Demonstration

Before the actual experiment, a demonstration program introduced subjects to the phenomenon of the missing fundamental. First, the single harmonics of a complex tone were presented in descending order. The sequence started with the 17th harmonic and finished with the 2nd. A complex tone was then generated by successively adding these overtones in descending order. The subjects were instructed to concentrate on a constant, low, buzzing tone, first perceptible approximately at the addition of the 11th harmonic and rising as supplementary harmonics were added. On a monitor, the participants were able to watch the harmonics to which they were being exposed. Finally, the subjects compared the complex tone without the fundamental with two sinusoidal tones. The frequencies of the sinusoids corresponded to those of the missing fundamental and second harmonic. Although the participants already knew that the second harmonic was the lowest component included in the complex tone, they were asked which sinusoid matched its perceived

pitch more closely. Thus they were directed toward the fact that they tended to refer to a pitch that was actually not present in the complex tone.

Fundamental Tuning Test

The distinctness of the virtual pitch impression was assessed by the accuracy with which its frequency could be determined. The subjects had to match the pitches of sinusoids with the perceived fundamental pitches of the corresponding overtone spectra by mouse control. Figure 2 shows a schematic diagram of a tuning task.

Calculation was made of both absolute and relative deviations of the tunings from the highest common submultiples of harmonics included in the complex tones. The deviations in hertz were then transformed to the relational cent scale. One cent is defined as one hundredth of a semitone, regardless of the absolute frequencies of the interval tones. The arithmetic mean of absolute average deviations of tunings from the reference frequency (expected value) was determined to provide a quantitative measure for the degree of task difficulty. (Some scientists, such as Lichte and Gray [1955], use the standard deviation of tunings from the reference frequency as opposed to the average deviation. All deviations are equally weighted for the average deviation, but large deviations are weighted to a higher degree than small ones for the standard deviation, as a consequence of a square term. The results of a study by Platt and Racine [1985], who calculated average deviations under different conditions, show that a task in which a sinusoid has to be tuned to a complex tone is easier than a task in which the sinusoid is used as the standard and the complex tone as the variable tone. Further results of the aforementioned study indicate that musicians perform significantly better than nonmusicians when they must tune a single tone to the perceived fundamental pitch of a complex tone consisting of the first eight harmonics.)

The optimum subjective frequency distance between the overtone spectrum and the corresponding missing fundamental was calculated via the relative average deviations of the tunings from the reference tone. In scientific terminology, negative deviations of the sinusoid from the expected value are referred to as "stretching,"

while positive deviations are referred to as "compression effects." (From musical practice and experiments, it is known that simultaneous as well as successive tone intervals are perceived as subjectively optimal when they are stretched—that is, when they are slightly larger than the mathematical ratio of the two tones would suggest [Fransson, Sundberg, & Tjernlund, 1974; Terhardt, 1969/70, 1971a, 1971b; Walliser, 1969; Ward, 1954]. Terhardt explains this effect by proposing that the complex tone properties [i.e., the intervals between their single partials] are first learned in early childhood and then later are used as a reference when listening to musical intervals. He refers to the observation that the pitch of a sinusoid is changed if another sinusoid is added. Both tones push the pitch of the other tone away from their own pitch and hence increase the subjectively perceived distance between them. On the other hand, Ohgushi [1983] reported that the octave enlargement phenomenon has its equivalent in neuronal discharge patterns. Ohgushi demonstrated by analysis of data from earlier experiments on squirrel monkeys [Rose, Brugge, Anderson, & Hind, 1967, 1968] that the peaks of the most frequently occurring neuronal interspike intervals are not exact submultiples of the applied frequency. The observed deviations correspond approximately to psychoacoustical data on octave stretching. Thus, the acquisition concept as well as the physiologically oriented model are able to explain the interval enlargement phenomenon. The effect of stretching is not only of importance for pure tone intervals, but also for the optimum subjective ratio between overtones and their fundamental.)

Stimulus Material

All harmonics of the complex tones were presented in cosine phase. The intensities of the harmonics were reduced in inverse proportion to their rank numbers. The timbres of such tones are more similar to those of several musical instruments than the sharp timbres of complex tones with harmonics of equal amplitude. The lowest presented harmonic within the spectrum was always set to an amplitude of 100% by the mixer of the synthesis program, regardless of its actual rank number n , to generate stimuli of similar loudness. The amplitudes of the higher harmonics were then decreased according to their rank numbers with respect to the fictitious intensity of the missing fundamental. Hence, the theoretical amplitude of the fundamental would always have amounted to a multiple of 100%, had it not been omitted.

Through this procedure, the possible influence of loudness differences on perceived pitch (Terhardt & Grubert, 1987), as it otherwise would have occurred between predominantly low and high spectra, was kept small. The weighted intensities of the complex tones, measured by a sound level meter (Bruel & Kjaer, Type 2232), varied between 60 and 70 dB(A), depending on the number of harmonics included. This range of variation was accepted, since Terhardt's model (Terhardt, Stoll, & Seewann, 1982a, 1982b) predicts a maximum subjective pitch difference of about 7 cents for the sinusoidal comparison tone when adjusted to two complex tone standards with an intensity difference of 10 dB SPL (Terhardt & Grubert, 1987). The intensity of the sinusoidal tone was 50 dB(A).

To suppress possible nonlinear distortion effects, the region around the missing fundamentals between 190 and 375 Hz was masked by digitally filtered, steep-slope bandpass noise with a weighted intensity of 35 dB. The applied noise level is sufficient to mask combination tones at intensities up to 30 dB (Zwicker, 1982).

The three spectral parameters of *number*, *relative height*, and *density of harmonics* were varied independently. By means of a computer program, for each of the eight spectral combinations (number [2] $\times$ height [2] $\times$ density [2]) seven possible overtone structures were calculated, yielding 56 different tasks. Each spectrum was randomly assigned a fictitious fundamental frequency between 200 and 365 Hz, which determined the actual frequencies of the applied harmonics.

Figure 2. Schematic diagram of a tuning task.

For the seven spectrally corresponding tasks, the average of absolute/relative deviations of tunings from their expected values in cents was calculated per subject. For further statistical analysis, this value was treated as one dependent measure. Each subject had to perform 112 evaluated tasks, which, divided by 7, yielded 16 values (number [2] $\times$ height [2] $\times$ density [2] $\times$ ear of presentation [2]). Left- and right-ear presentation differed only as far as the sequence of the 56 identical items was concerned.

Number of harmonics. The condition *few harmonics*, containing 3–6 overtones, was compared with the condition *many harmonics*, containing 7–10 overtones. This discrimination was made to evaluate the influence of spectral redundancy on the distinctness of virtual pitch given by different harmonics referring to the same fundamental. Schouten (1938, 1940a, 1940b) used stimuli with a large, indefinite number of components. De Boer (1956a) reduced the number to 7 and 5. Ritsma (1962) carried out experiments with 3 harmonics and finally Smoorenburg (1970) studied the perception of signals with only 2 frequency components. Under all these conditions, perception of virtual pitch was reported.

Height of harmonics in relation to missing fundamental (defined as the average rank number of the harmonics). The condition *low*, with an average of $n \leq 7$, was compared with *high*, with an average of $n > 7$. Low spectra mainly, but not exclusively, were built up of harmonics with rank numbers below 7, whereas high spectra predominantly consisted of harmonics with rank numbers above 7. The limit of $n = 7$ was chosen because the resolving power of the human ear for the distinction of single components extends to about the 7th harmonic (Plomp, Wagenaar, & Mimpen, 1973). Plomp (1967) and Ritsma (1967) found that the 3rd–6th harmonics are dominant for the formation of virtual pitch. Hence, the first condition included mainly well-resolved, dominant components, whereas the second condition comprised mainly high, unresolved harmonics. The highest included harmonics had a rank number of $n = 17$ and hence lay under the limit of n (lowest harmonic) = 20 for virtual pitch formation in three component signals with a missing fundamental of 200 Hz, as reported by Ritsma (1962).

Density of harmonics. The condition *narrow*, referring to sounds with only adjacent harmonics, was compared with the condition *wide*, referring to sounds containing one or more spectral gaps. A gap was generated by the omission of a single overtone within a harmonic series. Omissions beyond the highest and lowest applied harmonics were not considered.

The introduction of gaps was restricted by the following specifications: (1) The number of gaps within one complex tone should not exceed 4. (2) At least two adjacent harmonics should be present within the spectrum. (3) The relative height of the missing harmonics should match that of the present harmonics; hence, for low spectra, the average rank number n of the gaps should have been ≤ 7 , and for high spectra > 7 . (4) The sum of gaps within the seven corresponding tasks should be equal for the four combinations of the variables *number* and *height*. Therefore, on the average, the number of gaps was not varied in proportion to the number of applied harmonics, but was kept constant.

The 56 tasks, listed with respect to their spectral characteristics, are specified in Table 1. The fictitious fundamental frequencies represent the expected virtual pitches of the complex tone stimuli. The range of possible tunings was always restricted to 400 cents, corresponding to the musical interval of a major third. That small, limited region was chosen to avoid additional tuning peaks at frequencies of adjacent higher harmonics. To exclude learning effects, the required pitches of the missing fundamentals, which lay between 200 and 365 Hz, were systematically distributed within the tuning range. Independently, the starting frequencies of the sinusoids were varied in steps of 6 cents within the 400-cent span and assigned to the specific tasks randomly. In Table 1, for the complex tone

spectra, the presence of Harmonics 2–17 is indicated by “+.” Gaps between the applied harmonics are labeled “G.”

Procedure

The participants were allowed to switch between the sinusoidal tone and the standard as often as was necessary. The pitch of the sinusoid was controlled by moving the mouse up and down. The mouse offered two different response-sensitivity settings. The subjects were instructed first to find a rough approximation and then to continue with fine tuning.

Response time was not limited. Performance results were not reported to the subjects, since it had been found in earlier studies that no relevant improvement in tuning performance was achieved as a result of feedback (Campbell & Small, 1963; Platt & Racine, 1985; Spiegel & Watson, 1984). Before each session, the participants had the possibility of familiarizing themselves with the experimental situation by performing five practice tasks, which are not included in the following statistical analysis. The subjects were permitted to repeat these practice runs as often as was necessary.

According to their overall absolute tuning scores, subjects with an average deviation above the median of 63 cents were classified as *poor* performers, and subjects with an average deviation of less than 63 cents were classified as *good* performers. If the score was lower than 25% or higher than 75% (with reference to the whole sample), subjects were classified as *extremely poor* or *extremely good* performers, respectively.

In addition, it was established whether the subjects found the tasks easy or difficult and whether they had applied a spontaneous or a deliberate strategy.

RESULTS

A repeated measures analysis of variance was performed, with number, relative height, density of harmonics, and ear of presentation as within-subjects factors, and musician/nonmusician as the between-subjects factor. After allocation of subjects to the groups of good and poor performers, a second ANOVA was performed with the between-subjects factor of overall tuning performance. In addition, *t* tests and chi-square procedures were performed. A summary of the results is given in Table 2.

Main Effects

Spectral parameters. The results concerning absolute deviations are represented graphically in Figure 3. Those concerning relative deviations are shown in Figure 4. Means are indicated by large circles and standard deviations by small squares. For the following descriptions, *z* values indicate the magnitude of relative deviations from the expected values. The tail probabilities of the *z* values and the percentages of stretching/compression effects are given in turn. For number of harmonics, with absolute average deviations, the comparison between spectra containing few (3–6) and many (4–10) harmonics revealed that performance scores were significantly better when many harmonics were involved [$F(1,114) = 21.8, p < .001$; $M(\text{few}) = 77$ cents, $M(\text{many}) = 68$ cents]. For relative average deviations, spectra with few harmonics provoked too low and spectra with many harmonics provoked too sharp tunings [$F(1,114) = 87.7, p < .001$; $M(\text{few}) =$

Table 1
Spectral Composition of the 56 Applied Complex Tone Stimuli

Spectrum of Harmonics	Sound No.	FFF (in Hz)	Tuning Range of Sinusoid (in Hz)	Starting Frequency of Sinusoid (in Hz)	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17
Few/low/narrow	1	365	312-394	385					+	+	+									
	2	233	218-275	232						+	+	+								
	3	263	231-292	268				+	+	+	+									
	4	266	262-330	320					+	+	+	+	+							
	5	326	264-333	289			+	+	+	+	+	+								
	6	362	336-423	367				+	+	+	+	+	+							
	7	209	187-236	225	+	+	+	+	+	+	+									
Few/low/wide	1	230	205-258	239					+	G	+	+								
	2	239	232-292	240	+	G	G	+	+	G	+									
	3	341	295-372	317			+	+	G	+	G	+	+							
	4	332	275-347	334			+	+	+	G	G	+	+							
	5	242	226-285	258			+	G	G	+	+	+	G	+						
	6	206	182-229	209	+	+	G	+												
	7	359	299-377	313			+	+	G	+	G	+	G	+	+					
Few/high/narrow	1	353	306-386	324								+	+	+						
	2	272	231-291	241										+	+	+	+	+	+	
	3	227	222-279	263											+	+	+			
	4	338	290-366	310							+	+	+	+	+					
	5	302	278-350	316													+	+	+	+
	6	305	299-377	355						+	+	+	+	+	+					
	7	257	248-312	291									+	+	+	+	+			
Few/high/wide	1	224	183-231	202									+	G	+	+				
	2	329	289-364	348										+	+	+	+	G	+	
	3	296	241-304	266									+	+	G	G	+	+	+	
	4	317	277-349	299							+	+	G	+	G	+				
	5	200	187-238	196				+	+	+	G	+	G	G	+	+				
	6	335	319-402	360							+	G	G	+	+		G	+		
	7	293	210-311	256												+	+	G	+	
Many/low/narrow	1	290	270-340	276			+	+	+	+	+	+	+							
	2	356	289-365	346	+	+	+	+	+	+	+	+	+							
	3	347	286-360	312		+	+	+	+	+	+	+	+							
	4	212	177-223	205		+	+	+	+	+	+	+	+							
	5	236	214-269	232	+	+	+	+	+	+	+	+								
	6	308	296-372	363	+	+	+	+	+	+	+	+	+	+						
	7	275	221-279	247		+	+	+	+	+	+	+	+	+						
Many/low/wide	1	203	167-210	191	+	+	+	G	+	G	G	+	+	+						
	2	278	252-318	266	+	+	+	+	G	+	G	+	+	+						
	3	269	229-289	277	+	+	+	+	+	G	+	+	+	+	+					
	4	251	219-275	252		+	+	+	G	+	+	+	+							
	5	320	294-370	327	+	+	+	+	G	G	+	+	G	+						
	6	311	297-375	333	+	+	G	+	+	G	+	+	G	+	+					
	7	221	204-257	219	+	+	+	+	G	+	G	+	+	+	+	+				
Many/high/narrow	1	245	230-290	258					+	+	+	+	+	+	+	+	+			
	2	299	241-304	284						+	+	+	+	+	+	+	+	+	+	
	3	344	327-412	402							+	+	+	+	+	+	+	+	+	
	4	215	212-267	220				+	+	+	+	+	+	+	+	+	+			
	5	323	294-371	351										+	+	+	+	+	+	+
	6	314	303-382	328									+	+	+	+	+	+	+	
	7	248	209-264	237						+	+	+	+	+	+	+	+			
Many/high/wide	1	218	214-252	220					+	G	+	+	G	+	+	+	+	+	+	
	2	254	225-284	237						+	G	+	+	G	+	+	+	+	+	
	3	350	312-394	364						+	+	G	G	+	+	+	+	+		
	4	287	272-342	281				+	+	+	+	+	G	G	+	+	G	+	G	+
	5	260	234-295	241				+	+	+	+	G	+	+	+	+	+	G	+	
	6	284	260-327	266							+	+	G	+	+	+	+	+	+	
	7	281	242-305	285					+	G	+	+	+	+	+	+	G	+		

Note—The FFFs (fictitious fundamental frequencies) are the highest common divisors of the included harmonics. They represent the virtual pitches of the complex tones that are to be expected theoretically. All tuning ranges for the sinusoidal tones in hertz correspond to a span of 400 cents. The starting frequencies for tuning of the sinusoids were determined in steps of 6 cents and assigned to the specific tasks randomly. For the overtone spectra, the presence of Harmonics 2-17 is indicated by “+.” Gaps between the applied harmonics are labelled as “G.”

Table 2
Main Effects for the Variables of
Number, Height, and Density of Harmonics, and Ear of Presentation and Musicians/Nonmusicians

Main Effect	Condition 1	M	σ	Condition 2	M	σ	F(df)	p
Number, absolute	few	77	37	many	68	45	$F(1,114) = 21.8$	<.001
Number, relative	few	-7	36	many	16	52	$F(1,114) = 87.7$	<.001
Height, absolute	low	70	44	high	75	39	$F(1,114) = 5.6$	=.018
Height, relative	low	10	45	high	-1	44	$F(1,114) = 11.4$	=.001
Density, absolute	narrow	65	38	wide	79	45	$F(1,114) = 60.3$	<.001
Density, relative	narrow	-10	35	wide	19	54	$F(1,114) = 79.7$	<.001
Ear, absolute	left	73	42	right	73.5	40	$F(1,114) = 0.16$	.69*
Ear, relative	left	4	41	right	5	41	$F(1,114) = 0.14$	.7*
Musicians/nonmusicians, absolute	musicians	45	22	nonmusicians	100	38	$F(1,114) = 97.6$	<.001
Musicians/nonmusicians, relative	musicians	-13	25	nonmusicians	22	52	$F(1,114) = 28$	<.001

Note—Results are given for absolute and relative deviations. Means and standard deviations are given in cents. *n.s.

Figure 3. Absolute average deviations of fundamental tunings in cents for the conditions "few/many harmonics," "low/high harmonics," and "narrow/wide spectrum." (Note that the considered pitch range of 300 cents corresponds to three semitones, i.e. the musical interval of a minor third.)

Figure 4. Relative average deviations of fundamental tunings in cents for the conditions "few/many harmonics," "low/high harmonics," and "narrow/wide spectrum."

-7 cents, $z(\text{few}) = 1.9$, $p = .026$, stretching = 0.4%; $M(\text{many}) = +16$ cents, $z(\text{many}) = 4.3$, $p < .01$, compression = 0.93%].

For relative height of harmonics, with absolute average deviations, low spectra ($n \leq 7$) caused more exact tunings than did high spectra ($n > 7$) [$F(1,114) = 5.6$, $p = .018$; $M(\text{low}) = 70$ cents, $M(\text{high}) = 75$ cents]. With relative absolute deviations, low spectra caused too sharp fundamental assessments [$F(1,114) = 11.4$, $p = .001$; $M(\text{low}) = +10$ cents, $z(\text{low}) = 2.6$, $p = .004$, compression = 0.58%; $M(\text{high}) = -1$ cent, $z(\text{high}) = 0.2$, $p = .44$, compression = 0.06%].

For density of harmonics, with absolute average deviations, the results indicate that spectral density, which was neglected in earlier studies, has a strong influence on the perceptibility of virtual pitch. Narrow spectra led to more accurate tuning performances than did wide spectra [$F(1,114) = 60.3$, $p < .001$; $M(\text{narrow}) = 65$ cents, $M(\text{wide}) = 79$ cents]. With relative average deviations, narrow spectra resulted in too flat, and wide spectra in too sharp tunings [$F(1,114) = 79.7$, $p < .001$; $M(\text{narrow}) = -10$ cents, $z(\text{narrow}) = 2.7$, $p = .004$, stretching = 0.58%; $M(\text{wide}) = +19$ cents, $z(\text{wide}) = 16.8$, $p < .01$, compression = 1.08%].

Ear of presentation. No difference was obtained for absolute and relative average deviations between right- and left-ear presentation [absolute average deviations, $F(1,114) = 0.16$, $p = .69$; relative average deviations, $F(1,114) = 0.14$, $p = .7$].

Musicians/nonmusicians. The results concerning absolute deviations are represented graphically in Figure 5. Those concerning relative deviations are shown in Figure 6.

Musicians [absolute average deviations, $F(1,114) = 97.6$, $p < .001$] achieved by far more accurate tuning results than did nonmusicians [absolute average deviations: $F(1,114) = 97.6$, $p < .001$; $M(\text{musicians}) = 45$ cents, $M(\text{nonmusicians}) = 100$ cents].

Figure 7 is a visual representation of the distribution of all raw data of each person as relative deviations from the expected values for musicians (a) and nonmusicians

Figure 5. Absolute average deviations of fundamental tunings in cents for musicians and nonmusicians.

(b); 6,496 values, derived by 112 tasks * 58 subjects per graph. The abscissa represents the scale for the negative/positive deviations in cents. The ordinate refers to the number of overall tunings (note that the scale is standardized by the peak of the distribution as a whole). Musicians tended to tune the fundamental too flat. The opposite effect is observed in nonmusicians. Among non-musicians, the flattening of the peak is also recognizable, but when tunings are averaged, this is compensated by the trend to tune the fundamental too sharp. No such tendency is observed among nonmusicians.

Statistical analysis reveals that both the stretching effect [$M(\text{musicians}) = -13$ cents, $z = 4.25$, $p < .01$, stretching = 0.72%] and the compression effect [$M(\text{non-musicians}) = +22$ cents, $z = 3.7$, $p < .01$, compression = 1.26%] are highly significant [$F(1,114) = 28$, $p < .001$].

Interactions

For an overview, all possible two-way interactions for absolute and relative deviations concerning the three spec-

Figure 6. Relative average deviations of fundamental tunings in cents for musicians and nonmusicians.

tral parameters and the variable of musicians/nonmusicians are represented in Table 3.

Spectral parameters. The interaction of number $\times$ height of harmonics [absolute average deviations, overall, $F(1,114) = 11$, $p = .001$] shows that the relative height of the harmonics influenced tuning performance only when the tone contained few spectral components. Under that condition, the predominance of low harmonics made fundamental assessment easier [$F(1,114) = 45.3$, $p < .001$; $M(\text{few-low}) = 70$ cents, $M(\text{few-high}) = 84$ cents]. On the other hand, the number of harmonics was important only when high components were dominant [$F(1,114) = 50$, $p < .001$; $M(\text{few-high}) = 84$ cents, $M(\text{many-high}) = 65$ cents]. When the tone included primarily low harmonics, a high number of harmonics did not result in an additional improvement in fundamental pitch perception. Obviously, each of the stimulus characteristics *many* and *low* harmonics is sufficient on its own to cause a strong fundamental pitch sensation. The task was not made easier by combining both conditions. No significant interactions were obtained for relative deviations.

The interaction of number $\times$ density of harmonics [absolute average deviations; overall, $F(1,114) = 24.7$, $p < .001$] shows that narrow harmonics led to a significantly better tuning performance only if many components were involved [$F(1,114) = 89.3$, $p < .001$; $M(\text{few-narrow}) = 76$ cents, $M(\text{many-narrow}) = 57$ cents]. This result reflects the fact that each adjacency between two harmonics gave an unequivocal indication of the fundamental pitch. With an increasing number of harmonics, information about the missing fundamental became more and more redundant. Spectral continuity caused a strong improvement of tuning results, especially when many harmonics contributed to the pitch sensation.

The interaction [relative average deviations; overall, $F(1,114) = 51.9$, $p < .001$] also shows that a tendency toward sharpened tunings was present in the many harmonics condition only if those harmonics included spectral gaps [$F(1,114) = 76.8$, $p < .001$; $M(\text{many-narrow}) = -8$ cents, $z = 2.2$, $p = .03$, stretching = 0.46%; $M(\text{many-wide}) = +40$ cents, $z = 11.1$, $p < .01$, compression = 2.24%]. This result supports the hypothesis that too sharp fundamental assessments are caused by spectra that contain dominant harmonics in equivocal configurations.

The interaction of height $\times$ density of harmonics [absolute average deviations; overall, $F(1,114) = 26.4$, $p < .001$] indicates that spectral density is an important factor only if the harmonics concerned lie in the low frequency region [$F(1,114) = 47$, $p < .001$; $M(\text{low-narrow}) = 58$ cents, $M(\text{low-wide}) = 84$ cents]. This result is considered to be a consequence of the principle of dominance. Since low harmonics contributed to a higher degree to the final pitch sensation than high harmonics, gaps in high frequency regions did not significantly impair the tuning performance.

The interaction [relative average deviations; overall, $F(1,114) = 61.9$, $p < .001$] also clearly demonstrates

Figure 7. Distribution of raw data derived from all tasks ($N=112$) and all subjects ($N=58$) in musicians (a) and nonmusicians (b).

that the tendency to tune the sinusoid too sharp in the low harmonics condition was relevant only for wide spectra [$F(1,114) = 83.5$, $p < .001$; $M(\text{low-narrow}) = -16$ cents, $z = 4.4$, $p < .01$, stretching = 0.92%; $M(\text{low-wide}) = +35$ cents, $z = 9.7$, $p < .01$, compression = 2.04%]. Therefore, this result further supports the assumption that difficult sound spectra with equivocal indications to the missing fundamental are responsible for the tendency to tune a sinusoid too sharply.

Spectral parameters $\times$ musician/nonmusician. The interaction of number of harmonics $\times$ musicians/nonmusicians [absolute average deviations; overall, $F(1,114) = 6$, $p = .017$] shows that the number of harmonics affected the tuning performance only of musicians to any significant degree [$F(1,114) = 25.3$, $p < .001$; $M(\text{musicians-few}) = 51$ cents, $M(\text{musicians-many}) = 38$ cents]. Among nonmusicians, this tendency was insignificant. Obviously, musicians benefited more from the informa-

Table 3
Two-Way Interactions for All Possible Combinations of the Variables of Number, Height, Density of Harmonics, and Musicians/Nonmusicians

Interaction	Condition 1	M	σ	Condition 2	M	σ	Condition 3	M	σ	Condition 4	M	σ	F(df)	p
Number $\times$ height, absolute	few-low	70	40	few-high	84	37	many-low	70	56	many-high	65	44	$F(1,114) = 11$	$= .001$
Number $\times$ height, relative	few-low	-.5	30	few-high	-13	40	many-low	20	62	many-high	12	48	$F(1,114) = 1.16$	$= .28^*$
Number $\times$ density, absolute	few-narrow	76	41	few-wide	79	36	many-narrow	57	37	many-wide	81	59	$F(1,114) = 24.7$	$< .001$
Number $\times$ density, relative	few-narrow	-11	31	few-wide	-3	40	many-narrow	-8	39	many-wide	40	73	$F(1,114) = 51.9$	$< .001$
Height $\times$ density, absolute	low-narrow	58	40	low-wide	84	55	high-narrow	74	37	high-wide	77	42	$F(1,114) = 26.4$	$< .001$
Height $\times$ density, relative	low-narrow	-16	35	low-wide	35	66	high-narrow	-3	38	high-wide	2	51	$F(1,114) = 61.9$	$< .001$
Musician $\times$ number, absolute	mus.-few	51	22	mus.-many	38	23	nonmus.-few	102	32	nonmus.-many	98	43	$F(1,114) = 6$	$= .017$
Musician $\times$ number, relative	mus.-few	-19	22	mus.-many	5	29	nonmus.-few	-7	43	nonmus.-many	39	61	$F(1,114) = 18.8$	$< .001$
Musician $\times$ height, absolute	mus.-low	42	24	mus.-high	49	22	nonmus.-low	99	41	nonmus.-high	101	35	$F(1,114) = 2.03$	$= .15^*$
Musician $\times$ height, relative	mus.-low	-5	28	mus.-high	-20	22	nonmus.-low	24	54	nonmus.-high	18	52	$F(1,114) = 3.6$	$= .057^*$
Musician $\times$ density, absolute	mus.-narrow	40	19	mus.-wide	49	25	nonmus.-narrow	90	36	nonmus.-wide	110	40	$F(1,114) = 8.4$	$= .005$
Musician $\times$ density, relative	mus.-narrow	-17.5	22	mus.-wide	-2	27	nonmus.-narrow	-8	43	nonmus.-wide	45	62	$F(1,114) = 35.3$	$< .001$

Note—Interactions are given for absolute and relative deviations. Means and standard deviations are given in cents. *n.s.

tion provided by additional harmonics than did nonmusicians. This result might have been due to the greater experience of musicians with different kinds of complex tones, which allowed a more exact identification of the stimulus material on the basis of its spectral composition. Therefore, clues to the fundamental pitch could be exploited to a greater extent than they were by untrained subjects.

The interaction also shows [relative average deviations; overall, $F(1,114) = 18.8, p < .001$] that the compression effect towards the spectrally present harmonics in the many harmonics condition occurred only in nonmusicians [$F(1,114) = 31.8, p < .001$; $M(\text{musicians-many}) = -7$ cents, $z = 1.9, p = .06$, stretching = 0.4%; $M(\text{nonmusicians-many}) = +39$ cents, $z = 10.8, p < .01$, compression = 2.28%].

The attractor effect of a distinct overtone spectrum was also present among musicians, but it was not strong enough to compensate for the stretching effect, which was found generally in this group. These findings support the conclusion that compression is a function both of the subjects' musical experience and of the distinctness of the overtone spectrum. The tendency to perceive the complex tone as slightly heightened increased with the difficulty of the task as perceived by the listener and with the prominence of the harmonics.

For the interaction of density of harmonics $\times$ musicians/nonmusicians [absolute average deviations; overall, $F(1,114) = 8.4, p = .005$], although narrow spectra significantly improved tuning performance both among musicians and nonmusicians, the effect is more pronounced in nonmusicians [$F(1,114) = 11.8, p = .001$; $M(\text{musicians-narrow}) = 40$ cents, $M(\text{musicians-wide}) = 49$ cents; $F(1,114) = 56.86, p < .001$; $M(\text{nonmusicians-narrow}) = 90$ cents, $M(\text{nonmusicians-wide}) = 110$ cents]. This effect might have been caused by the greater experience of musicians in dealing with different complex tone spectra and in realizing their fundamental pitch even when parts of their complete patterns were missing. Probably, basic gestalt attributes of incomplete spectra were still available to musicians under aggravated conditions, whereas performance decreased more rapidly in nonmusicians.

The results [relative average deviations; overall, $F(1,114) = 35.3, p < .001$] also indicate that wide spectra only result in sharpened tunings among nonmusicians [$F(1,114) = 41.2, p < .001$; $M(\text{musicians-wide}) = -8$ cents, $z = 2.2, p = .03$, stretching = 0.46%; $M(\text{nonmusicians-wide}) = +45$ cents, $z = 12.5, p < .01$, compression = 2.63%]. Although narrow spectra caused a flatter pitch sensation among musicians than did wide spectra, the tunings exhibited stretching effects under both conditions. The interaction can again be interpreted as a function of the coincidence of two factors that impair fundamental pitch assessment—namely, incomplete, ambiguous spectra and the lack of musical experience. Under this difficult condition, pitch perception exhibited a tendency toward compression—that is, toward the higher harmonics that were physically present in the tone.

Objective performance rating/subjective assessment.

A strong relationship was observed between the allocation of subjects to the groups of poor and good performers and their subjective assessment of task difficulty [$\chi^2(1) = 18.6, p < .001$]. This coincidence becomes even clearer when one compares the extreme groups [$\chi^2(1) = 19.2, p < .001$]. The relative deviations show that the good performers tended to tune the fundamental too flat, whereas the opposite was true of the poor performers [$t = 5.6, p < .01$; $M(\text{good}) = -14$ cents, $z(\text{good}) = 4, p < .01$, stretching = 0.82%; $M(\text{poor}) = +9.5$ cents, $z(\text{poor}) = 2.6, p = .01$, compression = 0.55%].

As expected, the majority of musicians belonged to the group of good performers, whereas the opposite was true for nonmusicians [$\chi^2(1) = 40, p < .001$]. Again, the allocation becomes even clearer when one compares only extreme groups [$\chi^2(1) = 50.1, p < .001$]. Furthermore, the extremely poor performers predominantly reported that they had applied a spontaneous strategy, whereas the extremely good performers preferred a deliberate approach [$\chi^2(1) = 3.6, p = .054$].

Moreover, musicians tuned the fundamental too flat, regardless of their overall performance, whereas nonmusicians tuned the fundamental too high only when they were poor performers [overall, $F(1,112) = 9.9, p = .002$]. In good performers, the magnitude of the stretching effect corresponded roughly to that of musicians [$M(\text{good}) = -16$ cents, $z = 4.4, p < .01$; stretching = 0.91%]. This result indicates that the stretching effect is not a phenomenon that can be attributed to musical practice (e.g., from the sound spectra of piano strings, which are composed of harmonics that are slightly stretched), but instead matches the ability of fundamental pitch perception in general.

DISCUSSION

Absolute Deviations

As expected, a high number of harmonics makes fundamental pitch perception easier. This result can be explained by the greater extent of unequivocal indications of the missing fundamental, derived from a higher amount of neuronal activation.

This result contrasts with Wiesmann and Fastl's (1992) finding that the distinctness of the virtual pitch impression is diminished by an increase in the number of harmonics involved. Since their experiment was based on the subjective evaluation of the salience of pitch, the result might be due to a confusion of spectral and virtual pitch impressions. I suggest that an increasing prominence of higher harmonics shifted subjects' attention towards spectral components. Therefore, it is possible that a weighting process in favor of the overtone spectrum resulted in the subjective impression of reduced significance of the fundamental, in spite of an actual increase in fundamental pitch distinctness.

The importance of the region of dominance for the formation of virtual pitch (Plomp, 1967; Ritsma, 1967;

Zatorre, 1987) was again demonstrated by the method of pitch matching. In line with previous studies, the low frequency region was shown to be essential for the perceptibility of the missing fundamental.

Density, and hence fundamental pitch ambiguity, was shown to have remarkable influence on tuning performance. But the highest amount of observed variation was due to differences among subjects with different degrees of musical experience. The possibility that the difference between the two groups is a consequence of unequal exposition to tuning in general can be discarded, because Platt and Racine (1985) reported that musically inexperienced subjects remained inferior to experienced subjects even after six extensive tuning sessions.

The high degree of interindividual variation suggests that virtual pitch formation is due to a highly flexible mechanism that is sensitive to learning throughout life. Although the possibility of genetic differences cannot be excluded (to judge from the present data), this explanation seems rather insufficient to account for the entire effect.

The result supports the view that virtual pitch formation takes place at a high level of neuronal processing, and it seems likely that cortical mechanisms are involved in the ability. This assumption agrees with neurological studies of brain-lesioned patients whose ability at virtual pitch perception was severely impaired after removal of the right hemispheric primary auditory cortex (Zatorre, 1987). Concepts that refer to a subcortical time-coding mechanism working on the basis of quite rigid neurophysiological principles seem to be rather inadequate to account for the great extent of interindividual variation observed.

Developmental studies, as well as comparative studies among different species, indicate that fundamental pitch perception must reflect an essential function. On the other hand, this ability seems to be rather sensitive to modification throughout life. It therefore seems plausible that this ability is at least partly mediated by similar learning processes in all species. In humans, the sensitivity to acoustical pattern acquisition is obviously not restricted to an early period of life. Therefore, musical education may strongly reinforce a primarily natural learning mechanism.

Relative Deviations

The results regarding relative deviations indicate that the stretching effect is a basic feature of complex tone perception. This finding agrees with the psychoacoustical concept of central pitch formation (Terhardt, 1971a, 1971b, 1982; Terhardt et al., 1982a, 1982b). An important characteristic of Terhardt's model is the discrimination between the analytical, spectrally oriented and the holistic, fundamental-oriented modes of perception. Since the first harmonic very often does not dominate a complex tone spectrum in loudness, fundamental pitch is mainly determined by virtual pitch, which is derived through the analysis of higher components. The model

calculates weights both for the relative salience of the included partials and for the resulting possible fundamental frequencies. Platt and Racine (1985) already pointed out in their paper that Terhardt's model does not yield predictions for the weighting of the analytic, overtone-oriented mode in contrast to the holistic, fundamental-oriented mode. The present results provide evidence that task difficulty has a strong influence on subjectively perceived pitch.

As virtual pitch allocation becomes too difficult, the spectral clues of the overtones become more important. The perceived pitch then represents a compromise between both modes of perception, and it corresponds neither to the pitch of the missing fundamental nor to that of a single spectral component. Terhardt's model predicts a clear decision in favor of one mode of perception and hence does not allow an intermediate solution, whereas the present results indicate a compromise, which seems to be a product of the weighting procedure. This finding again provides evidence that the final pitch percept is strongly determined by subjective evaluation functions that depend on previous experience of the acoustical stimulus material. Pitch evaluation in favor of spectral or virtual aspects may take place at a high level of neuronal processing.

The ability of an overtone spectrum to attract the perceiver's attention depends on its prominence. Spectra including many harmonics from the low region of dominance give rise to the most distinct spectral impression. As a consequence, such overtone structures also facilitate fundamental pitch assessment. When gaps impair the perceptibility of the fundamental in a prominent spectrum, thus introducing a certain amount of virtual pitch ambiguity, the weighting process is supposed to favor the remaining spectral aspect of information. Accordingly, when the spectrum is less prominent, either because of only few or predominantly high harmonics, gaps do not evoke compression effects.

The same shift from the virtual aspect of information toward the spectral aspect is observed in nonmusicians and in poor performers, respectively. It may be that difficulty, caused either by stimulus characteristics or by the lack of musical abilities, is the common factor in both observed compression effects.

On the average, tones built up of many harmonics exhibit a tendency to be perceived as slightly sharper than tones comprising only few harmonics. But when interactions (number $\times$ density, number $\times$ musicians/nonmusicians) are analyzed further, it becomes evident that this effect is present only under aggravated conditions. As mentioned above, the spectral aspect of information dominates the virtual aspect in tones with a high number of harmonics when the task is sufficiently difficult.

The height effect can be interpreted as a consequence of the superposition of the included overtone intervals. A wide distance between the spectral harmonics and the missing fundamental suggests a larger amount of stretching than a small distance. This explanation holds both for the concept of learning optimal interval sizes (Terhardt,

1969/70, 1971a, 1971b) and for the neurophysiologically based approach of Ohgushi (1983).

The difference between the low and the high conditions corresponds well with the findings of Walliser (1969b), Terhardt (1971a, 1971b) and Lichte (1941), who observed that stretching effects between overtones and the perceived fundamental pitch depend on the relative height of the spectra.

The present results on relative deviations must also be discussed in light of the historical debate on the interrelation between the spectral composition of complex tones and their perceived pitch. In 1895, Helmholtz stated that "upper partial tones ... give a compound tone a brighter and higher effect." Banister (1934), in *A Handbook of General Experimental Psychology*, wrote that "the pitch of a complex tone is usually slightly higher than that of a pure tone of the same frequency," and hence suggested a compression effect between the actual and the perceived fundamental pitch. Lichte (1941) investigated how the presence or absence of high harmonics influences the pitch of the corresponding fundamental. In this experiment, subjects had to match the pitches of sinusoidal tones with complex tones with a fundamental pitch of 180, 360, or 540 Hz. At 180 Hz, the complex tones consisted of 16 harmonics. Tones with harmonics that were attenuated in ascending order by 1 dB each (low condition) were compared with tones with harmonics that were attenuated by 1 dB in descending order, beginning with the 15th harmonic and finishing with the fundamental (high condition). At 360 Hz, the first 8 harmonics were utilized; at 540 Hz, only the first 5 harmonics were presented. Six nonmusicians participated in the study. At 360 and 540 Hz, the results revealed that predominantly high spectra were perceived as significantly flatter than predominantly low spectra. At 180 Hz, the same tendency was of no significance. This result agrees well with the present data.

In 1955, Lichte and Gray tried to replicate Lichte's (1941) previous results with a similar experimental design. Harmonic complex tones with fundamental frequencies of 250 and 700 Hz were low-pass filtered at different slopes. For the flat-slope filter condition, the 250-Hz tone consisted of more than 25 harmonics and the 700-Hz tone included 19 harmonics. For the steep-slope filter condition, the 250-Hz tone consisted of the first 10 harmonics, and the 700-Hz tone included the first 5 harmonics. Only individuals with sufficient scores on the Seashore Measures of Musical Talents were admitted to participate in the experiments. The results indicated effects exactly opposite to those found in 1941. At both frequencies, complex tones with mainly high harmonics were judged to be sharper in pitch than tones with mainly low harmonics. The difference amounted to 52 cents at 250 Hz and 24 cents at 700 Hz. Unfortunately, these results cannot be compared directly with the present ones. Lichte used the standard deviation of the sinusoid of the first harmonic of the complex tone as a measure for the perceived pitch, whereas in the present study, the arithmetic mean of the absolute deviations was calculated from the ficti-

tious missing fundamental. This fact may account for the relatively high deviation of 52 cents at 250 Hz, in comparison with the deviation of 10.1 cents measured in the present study for a comparable frequency range.

The present results suggest a possible interpretation of the observed discrepancies. In Lichte's (1941) first study, the number of harmonics for both the low and the high spectral conditions was kept constant, but this was not true in his second study. In the latter experiment, the parameters of height and number of harmonics were varied simultaneously, thus probably leading to a confusion of both effects. Because the ratios of the numbers of harmonics in Lichte's low and high conditions amounted to 10/25 for the 250-Hz tone and 5/19 for the 700-Hz tone, it is to be expected that the number of harmonics involved had a strong influence on pitch assessments. Hence, Lichte's controversial findings might be explained by a predominance of the number effect over the height effect.

The data from the present experiment reveal that the involvement of predominantly high harmonics results in a slightly flattened pitch sensation. The presence of a high number of harmonics shifts the subjective pitch slightly upward when the task is difficult for the listener, either because of a lack of musical experience or because of incomplete, ambiguous spectra. Greer (1970) reported that brass players tend to intonate too flat by 20 cents when trying to adjust the pitch of their instrument to a sinusoid. No such compression effect was observed when the reference tone was given by another musical instrument. Obviously, subjects paid more attention to spectral clues than to virtual pitch. It can only be speculated that this bias was due to the incompleteness of the applied sound spectrum or to some other influence aggravating fundamental pitch perception.

Because the results of Lichte's second study (Lichte & Gray, 1955) contradicted psychoacoustical models concerning central pitch perception (Goldstein, 1973; Terhardt, 1971a; Wightman, 1973), Terhardt (1971b) carried out a similar tuning experiment which confirmed his own theoretical predictions. A stretching effect between the actual fundamental and the perceived virtual pitch of the complex tone was found.

Platt and Racine (1985) found that a sine wave was tuned about 10.2 cents too sharp when adjusted to a complex tone stimulus consisting of the first eight harmonics. Thus, their results were again in contradiction to theoretical concepts. Terhardt and Grubert (1987) attributed these discrepancies to differences in the applied sound pressure levels. Yet the authors stated that the observed discrepancies could not be attributed fully to intensity differences.

Given the present results, one must ask whether the stimuli used by Platt and Racine (1985) were too difficult for the subjects who participated in their experiment. Indeed it is evident from Platt and Racine's data that non-tuners (nonmusicians and subjects playing an instrument with fixed tuning) exhibiting larger absolute deviations showed a stronger trend to tune the sinusoid too sharp

than did tuners (subjects playing an instrument that has to be tuned by the musician) do. A comparison between the two applied experimental conditions gives additional support to that explanation. The situation in which the complex tone served as the standard and the sinusoid as the comparison tone was compared to the reverse situation. Absolute deviations revealed that the first task was much easier for both groups than the second task. Interestingly, this easier condition also resulted in a smaller compression effect than the more difficult condition did. In agreement with the findings of the present study, the largest compression effect was observed among nontuners on the more difficult task.

Hence, it can be speculated that the applied sound pressure levels as well as the general task difficulty affected Platt and Racine's (1985) data. The coincidence of both factors might explain, at least in part, the discrepancies between their findings and theoretical predictions. Thus, it can be concluded that theoretical models yield good predictions for virtual pitch, provided that the perceiving subjects are sufficiently able to extract the missing fundamental from the given spectral clues.

REFERENCES

- ANDERSON, D. J., ROSE, J. E., HIND, J. E., & BRUGGE, J. F. (1971). Temporal position of discharges in auditory nerve fibers within the cycle of a sine-wave stimulus: Frequency and intensity effects. *Journal of the Acoustical Society of America*, **9**, 1131-1139.
- BANISTER, H. (1934). Auditory phenomena and their stimulus correlations. *Handbook of general experimental psychology* (C. Murchison, Ed.) Worcester, MA: Clark University Press.
- BUNDY, R. S., & COLOMBO, J. (1982). Pitch perception in young infants. *Developmental Psychology*, **18**, 10-14.
- CAMPBELL, R. A., & SMALL, A. M. (1963). Effect of practice and feedback on frequency discrimination. *Journal of the Acoustical Society of America*, **35**, 1511-1514.
- CLARKSON, M. G., & CLIFTON, R. K. (1984). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. *Journal of the Acoustical Society of America*, **77**, 1521-1528.
- CYNN, J., & SHAPIRO, M. (1986). Perception of the missing fundamental by a species of songbird, the European starling (*Sturnus vulgaris*). *Journal of Comparative Psychology*, **100**, 356-360.
- DE BOER, E. (1956a). Pitch of inharmonic signals. *Nature*, **178**, 535-536.
- DE BOER, E. (1956b). *On the residue in hearing*. Unpublished doctoral dissertation, University of Amsterdam.
- EDDINGTON, D. K. (1980). Speech discrimination in deaf subjects with cochlear implants. *Journal of the Acoustical Society of America*, **68**, 885-891.
- FLETCHER, H. (1929). *Speech and hearing*. London: Macmillan.
- FLETCHER, H. (1931). Auditory patterns. *Review of Modern Physics*, **3**, 258-278.
- FRANSSON, F., SUNDBERG, J., & TJERNLUND, P. (1974). The scale of music. *Swedish Journal of Musicology*, **56**, 49-54.
- 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.
- GREEK, R. D. (1970). The effect of timbre on bass-wind intonation. In E. Gordon (Ed.), *Experimental research in the psychology of music* (pp. 65-94). Iowa City: University of Iowa Press.
- GRUBER, J. (1992). Periodizitätstönhöhe: Ergebnis einer Zeitstrukturanalyse? In W. Heinicke (Ed.), *Proceedings of the DAGA* (pp. 1599-1600). Berlin: Physik-Verlag.
- HEFFNER, H., & WHITFIELD, I. C. (1976). Perception of the missing fundamental by cats. *Journal of the Acoustical Society of America*, **59**, 915-919.
- HOUTSMA, A. J. M., & GOLDSTEIN, J. L. (1971). *Perception of musical intervals: Evidence for the central origin of the pitch of complex tones* (Tech. Rep. No. 484). Cambridge: Massachusetts Institute of Technology, Research Laboratory of Electronics.
- HOUTSMA, A. J. M., & GOLDSTEIN, J. L. (1972). The central origin of the pitch of complex tones: Evidence from musical interval recognition. *Journal of the Acoustical Society of America*, **51**, 520-529.
- JAVEL, E. (1986). Basic response properties of auditory nerve fibers. In R. A. Altshuler (Ed.), *Neurobiology of hearing: The cochlea* (pp. 213-245). New York: Raven Press.
- KIANG, N. Y. S., WATANABE, T., THOMAS, E. C., & CLARK, L. F. (1965). *Discharge patterns of single fibers in cat's auditory nerve* (MIT Research Monographs, No. 34). Cambridge, MA: Technology Press.
- LANGNER, G. (1983). Evidence for neuronal periodicity detection in the auditory system of the guinea-fowl: Implications for pitch analysis in the time domain. *Experimental Brain Research*, **52**, 333-355.
- LANGNER, G., & SCHREINER, C. E. (1988). Periodicity coding in the inferior colliculus of the cat: 1. Neuronal mechanisms. *Journal of Neurophysiology*, **6**, 1799-1822.
- LICHTE, W. H. (1941). Attributes of complex tones. *Journal of Experimental Psychology*, **28**, 455-480.
- LICHTE, W. H., & GRAY, F. R. (1955). The influence of overtone structure on the pitch of complex tones. *Journal of Experimental Psychology*, **49**, 431-436.
- LICKLIDER, J. C. R. (1951). A duplex theory of pitch perception. *Experientia*, **7**, 128-134.
- MEDDIS, R., & HEWITT, M. J. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery: I. Pitch identification. *Journal of the Acoustical Society of America*, **89**, 2866-2882.
- MILLER, G. A., & TAYLOR, W. G. (1948). The perception of repeated bursts of noise. *Journal of the Acoustical Society of America*, **20**, 171-180.
- OHGUSHI, K. (1983). The origin of tonality and a possible explanation of the octave enlargement phenomenon. *Journal of the Acoustical Society of America*, **73**, 1694-1700.
- PATTERSON, R. D. (1969). Noise masking of a change in residue pitch. *Journal of the Acoustical Society of America*, **45**, 1520-1524.
- PFEIFFER, R. R., & MOLNAR, C. E. (1970). Cochlear nerve fiber discharge patterns: Relationship to the cochlear microphonic. *Science*, **167**, 1614-1616.
- 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. (1967). Pitch of complex tones. *Journal of the Acoustical Society of America*, **41**, 1526-1533.
- PLOMP, R., WAGENAAR, W. A., & MIMPEN, A. M. (1973). Musical interval recognition with simultaneous tones. *Acustica*, **29**, 101-109.
- RITSMA, R. J. (1962). Existence region of the tonal residue. *Journal of the Acoustical Society of America*, **34**, 1224-1229.
- RITSMA, R. J. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *Journal of the Acoustical Society of America*, **42**, 191-198.
- RITSMA, R., & ENGEL, F. (1964). Pitch of frequency-modulated signals. *Journal of the Acoustical Society of America*, **42**, 191-198.
- ROSE, J. E., BRUGGE, J. F., ANDERSON, D. J., & HIND, J. E. (1967). Phase locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, **30**, 769-793.
- ROSE, J. E., BRUGGE, J. F., ANDERSON, D. J., & HIND, J. E. (1968). Patterns of activity in single auditory nerve fibers of the squirrel monkey. In A. V. S. de Reuck & J. Knight (Eds.), *Hearing mechanisms in vertebrates* (pp. 144-168). London: Churchill.
- SCHOUTEN, J. F. (1938). The perception of subjective tones. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, **41**, 1086-1094.
- SCHOUTEN, J. F. (1940a). The residue and the mechanism of hearing. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, **43**, 991-999.
- SCHOUTEN, J. F. (1940b). The residue, a new component in subjective sound analysis. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, **43**, 356-365.
- SEDLMEIER, H. (1992). *Tonhöhenwahrnehmung beim falschen Vampir*

- Megaderma Lyra*. Unpublished doctoral dissertation, Department of Biology, Ludwig-Maximilians-Universität, Munich.
- SEEBECK, A. (1843). Über die Sirene. *Annals of Physical Chemistry*, **60**, 449-481.
- SMOORENBURG, G. F. (1970). Pitch perception of two frequency stimuli. *Journal of the Acoustical Society of America*, **48**, 924-942.
- SPIEGEL, M. F., & WATSON, C. S. (1984). Performance on frequency discrimination tasks by musicians and nonmusicians. *Journal of the Acoustical Society of America*, **76**, 1690-1695.
- STOLL, G. (1980). Psychoakustische Messungen der Spektraltonhöhenmuster von Vokalen. In *Fortschritte der Akustik* (pp. 631-634). Berlin: VDE-Verlag.
- TERHARDT, E. (1969-1970). Oktavspreizung und Tonhöhenverschiebung bei Sinustönen. *Acustica*, **22**, 345-351.
- TERHARDT, E. (1971a). Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon. In *Proceedings of the 7th International Congress on Acoustics* (Vol. 3, p. 621). Budapest: Akademiai Kiado.
- TERHARDT, E. (1971b). Die Tonhöhe harmonischer Klänge und das Oktavintervall. *Acustica*, **24**, 126-136.
- TERHARDT, E. (1972). Zur Tonhöhenwahrnehmung von Klängen II. Ein Funktionsschema. *Acustica*, **26**, 187-198.
- TERHARDT, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, **55**, 1061-1069.
- TERHARDT, E. (1976). Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz. *Acustica*, **36**, 121-137.
- TERHARDT, E. (1978). Psychoacoustic evaluation of musical sounds. *Perception & Psychophysics*, **23**, 483-492.
- TERHARDT, E. (1982). Die psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmische Bestimmung. In C. Dahlhaus & M. Krause (Eds.), *Tiefenstruktur der Musik* (pp. 23-50). Berlin: TU-Berlin.
- TERHARDT, E. (1989). Warum hören wir Sinustöne? *Naturwissenschaften*, **76**, 496-504.
- TERHARDT, E. (1991). Music perception and sensory information acquisition: Relationships and low-level analogies. *Music Perception*, **8**, 217-240.
- TERHARDT, E., & GRUBERT, A. (1987). Factors affecting pitch judgments as a function of spectral composition. *Perception & Psychophysics*, **42**, 511-514.
- TERHARDT, E., STOLL, G., & SEEWANN, M. (1982a). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, **71**, 679-688.
- TERHARDT, E., STOLL, G., & SEEWANN, M. (1982b). Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *Journal of the Acoustical Society of America*, **71**, 671-678.
- TOMLINSON, W., & SCHWARZ, D. (1988). Perception of the missing fundamental in nonhuman primates. *Journal of the Acoustical Society of America*, **84**, 560-565.
- WALLISER, K. (1969a). Über die Spreizung von empfundenen Intervallen gegenüber mathematisch-harmonischen Intervallen bei Sinustönen. *Frequenz*, **23**, 139-143.
- WALLISER, K. (1969b). Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe. *Acustica*, **21**, 319.
- WARD, W. D. (1954). Subjective musical pitch. *Journal of the Acoustical Society of America*, **26**, 369-380.
- WEVER, E. G. (1949). *Theory of hearing*. New York: Wiley.
- WIESMANN, N., & FASTL, H. (1992). Ausgeprägtheit der virtuellen Tonhöhe und Frequenzunterschiedsschwellen von harmonischen komplexen Tönen. In W. Heinicke (Ed.), *Proceedings of the DAGA* (pp. 1600-1601). Berlin: Physik-Verlag.
- WIGHTMAN, F. L. (1973). Pitch and stimulus fine structure. *Journal of the Acoustical Society of America*, **54**, 397-406.
- ZATORRE, R. J. (1987). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, **84**, 566-572.
- ZWICKER, E. (1982). *Psychoakustik*. Springer-Verlag.

(Manuscript received September 2, 1992;
revision accepted for publication April 8, 1993.)