On the perceptual limits of octave harmony and their origin

Laurent Demany\textsuperscript{a} and Catherine Semal

Laboratoire de Psychoacoustique, Université de Bordeaux 2, 146 rue Léo-Saignat, F-33076 Bordeaux, France

Robert P. Carlyon

Laboratory of Experimental Psychology, Sussex University, Brighton BN1 9QG, England

(Received 18 January 1991; accepted for publication 12 August 1991)

Three subjects were monaurally presented with dyads of simultaneous, frequency modulated, sine tones approximately 1 oct apart. The tones, with carrier frequencies $F_L$ and $F_H$, were presented in a background of pink noise and at a low sensation level, so that they were completely resolvable by the subjects' peripheral auditory filters. In two experiments, subjects had to judge on each trial which of two dyads was inharmonic ($F_H \neq 2F_L$); the relative mistuning of the inharmonic dyad $[(F_H - 2F_L)/2F_L]$ was varied independently of $F_L$ and could be either positive or negative; $F_L$ was fixed within trials in one experiment, and varied within trials (from about 8%-20%) in the other experiment. In both experiments, performance monotonically worsened as $F_L$ increased from 300-2000 Hz; in addition, negative mistunings were more often identified as inharmonicities than positive mistunings. In another experiment, a 41-2AFC procedure was used to assess the detectability of changes in $F_H$, irrespective of their effect on the perception of harmonicity. Compared to the other two experiments, performance was not the same function of $F_L$, and was much better at high $F_L$. Thus it is concluded that the detection of inharmonicity was limited by factors other than those responsible for fine pitch discriminations. Our results also have implications with respect to the origin of human listeners' melodic octave templates.

PACS numbers: 43.66.Hg, 43.75.Bc [WAY]

\textsuperscript{a} Also at Laboratoire d'Audiologie Expérimentale, INSERM U.229, Hôpital Pellegrin, F-33076 Bordeaux, France.

\section*{INTRODUCTION}

Normal adult listeners, including quite amusical people, possess some ability to discriminate inharmonic from harmonic frequency relationships; this perceptual ability has been investigated in a number of experiments, using very different methods and stimuli (Broadbent and Ladefoged, 1957; Rasch, 1978; Moore \textit{et al.}, 1985b, 1986; Beerends and Houtsma, 1989; Hartmann, 1988; Hartmann \textit{et al.}, 1990; Demany and Semal, 1988, 1989, 1990; Carlyon and Stubbs, 1989; Carlyon, 1991; Carlyon \textit{et al.}, 1992).

As pointed out by Moore \textit{et al.} (1985b), there are basically two possible ways to detect inharmonicities. Consider a dyad of simultaneous pure tones with similar levels and different frequencies. If the levels and frequencies of the two tones are such that the tones stimulate common peripheral auditory filters, then the output waveform of these common filters will have a quite different period when the dyad is harmonic than when it is slightly inharmonic; the detection of inharmonicity can then rest on the detection of "beats of mistuned consonances" (Plomp, 1976, Chap. 3). However, inharmonicity detection is also possible in the absence of interactions between the peripheral representations of the tones (Demany and Semal, 1988, 1989; Carlyon, 1991; Carlyon \textit{et al.}, 1992). When simultaneous pure tones have harmonic relationships, they tend to be perceptually fused into a single sound entity, even if their cochlear representations do not interact; by contrast, inharmonic frequency relationships tend to promote perceptual segregation; thus, in the absence of peripheral interactions, inharmonicity detection can rest on fusion/segregation cues.

The present paper bears on the detection of inharmonicity for dyads of pure tones that do not interact in the cochlea. More specifically, we are concerned with the detection of deviations from octave relationships. In a previous study (Demany and Semal, 1990), two of us assessed the accuracy with which adult listeners are able to adjust two simultaneous and monaurally presented pure tones to be 1 oct apart. Each dyad was frequency modulated in an attempt to prevent the subjects from making sequential comparisons of its components' spectral pitches, i.e., from making melodic rather than harmonic judgments. The main independent variable was the carrier frequency of the lower tone ($F_L$) and subjects adjusted the carrier frequency of the higher tone. It was found that the accuracy of subjects' adjustments was much poorer for $F_L > 1000$-1500 Hz than for $F_L < 500$ Hz, if simultaneous octaves induced a clear-cut and distinctive sensation of harmony only for $F_L < 500$ Hz. Since melodic octaves, made up of successive pure tones, can be adjusted accurately when the lower tone is above 1000-1500 Hz (Ward, 1954; Demany and Semal, 1990), this result was rather surprising. However, it could be partly anticipated from a previous experiment in which octave harmony was studied with dichotic stimuli and a quite different method (Demany and Semal, 1988). It is also consistent with data reported by Hartmann \textit{et al.} (1990) on human listeners' ability to hear a mistuned harmonic out of an otherwise periodic complex tone with many spectral components; these authors...
found that above 2000 Hz, mistuned harmonics became very difficult to hear out. In the study reported here, our aim was to investigate again the perceptual limits of octave harmony, with a new method, and to get some insight on their origin.

There are two possible classes of explanation for the apparent deterioration of octave harmony when \( F_L > 1000 - 1500 \) Hz. A first hypothesis is that the central auditory system of human listeners does not contain equally accurate “harmonic octave templates” for the processing of low- and high-frequency complex tones. An alternative hypothesis is that the simultaneous presentation of two frequency components, even if resolved by the cochlea, degrades the internal representation of one or both of their frequencies when \( F_L \) is high. These two hypotheses are illustrated in Fig. 1.

According to the second hypothesis, the central “harmony recognizer” is equally effective when \( F_L \) is low or high, but in the latter case it receives noisy spectral information. Thus, at high frequencies, perceptual difficulties should occur not only with respect to harmony recognition: It should also be difficult to discriminate frequency changes in the individual components of complex tones, irrespective of the effect produced by these frequency changes on perceived harmony. In one of the experiments reported here, this prediction was tested on pure-tone dyads in the vicinity of the octave.

The second hypothesis was not a priori unlikely since it is known that the detection of sinusoidal modulations in amplitude or frequency can be degraded by the presentation of simultaneous frequency components that do not produce any direct masking of the signal (Wakefield and Viemeister, 1985; Yost et al., 1989; Wilson et al., 1990). Moore et al. (1984) previously investigated the discrimination of frequency changes in individual partials of periodic or quasi-periodic complex tones. However, they used spectrally rich tones at a single fundamental frequency, 200 Hz; in addition, the level of each partial was 60 dB SPL, too high to preclude within-channel interactions between the peripheral representations of even low-ranked partials. The pure-tone dyads used in each of the present experiments were presented in a noise background and at a low sensation level in order to ensure that their components would not interact in the cochlea.

Another goal of the present research was to determine if, for a given value of \( F_L \), the lower and upper frequencies corresponding to the limits of octave harmony are equally or inequally distant from \( 2F_L \). Many previous studies showed that a melodic octave, made up of successive pure tones, is generally perceived as well tuned for a range of frequency ratios centered significantly above 2/1 (see, e.g., Ward, 1954; Terhardt, 1970; Dobbins and Cuddy, 1982; Demany and Semal, 1990). Is there a similar phenomenon with respect to octave harmony? The adjustment experiments of Demany and Semal (1989, 1990) weakly suggest that such is the case. We used here a better procedure—a forced-choice task—to examine that question.

I. EXPERIMENT I

In this first experiment, the detectability of inharmonicity for dyads of pure tones in the vicinity of the octave was assessed as a function of the frequency of their lower component \( (F_L) \), which was varied from about 300 to 2000 Hz.

A. Method

1. Preliminary considerations

As pointed out by Moore et al. (1985b, 1986) and Hartmann (1988), experiments on the detection of inharmonicity have to avoid a number of methodological pitfalls in order to provide unambiguous results. Basically, one has to ensure that subjects base their judgments on harmony per se rather than pitch height or some other cue. One particular pitfall was not considered previously but had to be taken seriously here. Assume that during repeated blocks of trials, the stimulus set includes one fixed harmonic dyad and, say, ten inharmonic dyads, presented each ten times less often than the harmonic dyad. If the 11 dyads are clearly discriminable from each other, but induce very similar sensations of harmony, so that differential judgments of harmony per se
3. Stimuli

Each stimulus consisted of two simultaneous pure tones with carrier frequencies $F_L$ and $F_H$, which were sinusoidally frequency modulated at a 4-Hz rate, in a parallel manner (so that a constant frequency ratio, equal to $F_H/F_L$, was maintained during a given stimulus). All modulations had a zero-to-peak depth corresponding to 5% of the carrier frequency and started at a positive-going zero crossing. Each stimulus had a total duration of 500 ms (corresponding to exactly two modulation cycles), and was gated on and off with 10-ms linear amplitude ramps. Stimuli were presented monaurally, through a TDH 39 earphone, in a background of continuous pink noise. This noise had an overall level of 41.5 dB A and a 1/f power spectrum (Gardner, 1978). Each component tone had a 600-ms delay. Note that since the mistuning of the inharmonic dyad could be positive as well as negative, the position of this stimulus on a given trial could not be identified on the basis of relative pitch cues. Even the inexperienced subject (HP) understood the task immediately, in spite of the complete absence of feedback.

Four nominal values of $F_L$ were used: 300, 500, 1000, and 2000 Hz. However, in order to avoid repetitions of individual dyads (cf. Sec. I A 1), the exact value of $F_L$ on a given trial was selected at random within a logarithmic frequency range extending from 91% to 109% of one of its four nominal values. Successive stimuli with the same value of $F_L$ were low-pass filtered at 8 kHz (Frequency Devices 678, 80 dB/oct). The outputs of the two digital-to-analog converters were then mixed in a stereo amplifier (Marantz PM-55), whose output was checked with a spectrum analyzer (Hewlett Packard 3561A).

4. Procedure

The subjects were individually tested in a double-walled soundproof booth (Gisol), where they sat in front of a keyboard connected to the computer (Compaq 386/20) controlling the experiment. On each trial, they heard two successive stimuli with the same value of $F_L$. These two dyads were separated by a 500-ms pause. One of them was harmonic ($F_H = 2F_L$) and the other was inharmonic; in the latter case, the value of $F_H$ was obtained by multiplying $2F_L$ by 1 ± 0.005, 0.01, 0.02, 0.03, 0.04, 0.06, or 0.09. The inharmonic dyad was presented at random in the first or second observation interval and the subject had to judge which interval contained the more consonant dyad by pressing one of two labeled keys. No feedback was provided, even during the training sessions. Each response initiated the next trial after a 600-ms delay. Note that since the mistuning of the inharmonic dyad could be positive as well as negative, the position of this stimulus on a given trial could not be identified on the basis of relative pitch cues. Even the inexperienced subject (HP) understood the task immediately, in spite of the complete absence of feedback.

Four nominal values of $F_L$ were used: 300, 500, 1000, and 2000 Hz. However, in order to avoid repetitions of individual dyads (cf. Sec. I A 1), the exact value of $F_L$ on a given trial was selected at random within a logarithmic frequency range extending from 91% to 109% of one of its four nominal values. Since the mistuning of $F_H$ could take 14 different values (seven positive and seven negative values), $4 \times 14 = 56$ types of trials were run. Each type of trial was run once within blocks of 56 trials. In each block, the nominal value of $F_L$ varied in a sawtooth manner from trial to trial, starting at random from one of the four possible values (e.g., 500, 1000, 2000, 300, 500, etc.). The ordering of the 14 different mistunings used for a given nominal value of $F_L$ was selected at random, and independently for the four nominal values of $F_L$.

In a daily session, 10 to 40 blocks of trials were run, with pauses à volonté. Each subject was tested for a total of 160 blocks.

B. Results and discussion

Figure 2 shows the individual psychometric functions obtained, in each of the four frequency registers, for positive (open circles) and negative (closed circles) mistunings. The most obvious finding is the existence of asymmetries in the
detection of inharmonicity: Quite generally, negative mistunings were more often identified as inharmonicities than positive mistunings; this occurred in 11 cases out of 12 if one considers only the mistunings for which the psychometric functions cross the 75 percent correct level. The largest asymmetry was obtained in the case of subject HP for $F_L = 1000$ Hz. Here, it can be seen that the psychometric function for positive inharmonicites begins by going below the chance level (50 percent correct), and takes a positive slope only when the mistuning exceeds +3%; this, and the asymmetry per se, suggests that a physically harmonic octave made up of pure tones in a frequency ratio of exactly 2/1 is perceived as significantly less harmonic than a dyad of tones with a somewhat larger frequency ratio. Similar though less pronounced trends can be observed for LD and HP when $F_L = 500$ Hz.

However, it should be recalled that for the two stimuli presented on any trial, exactly the same $F_L$ value was used. Assume that on trials where the two stimuli could not be clearly differentiated from the point of view of harmony, the subjects tended to decide that the inharmonic stimulus was the stimulus with the lower virtual pitch (or the stimulus inducing the lowest spectral pitch). Such a strategy leads to the asymmetry actually obtained.

The potential influence of pitch biases was removed in Fig. 3, by averaging the data for positive and negative mistunings for each subject and frequency register. In each

![FIG 2](image-url) Psychometric functions for the detection of inharmonicity in experiment 1. Each column shows a given subject's data, in four panels corresponding to the four nominal values of $F_L$ used (from bottom to top: 300, 500, 1000, and 2000 Hz). In each panel, the parameter is the direction of mistuning. Open circles: positive mistunings ($F_n > 2F_L$); closed circles: negative mistunings ($F_n < 2F_L$).

![FIG 3](image-url) Psychometric functions for the detection of inharmonicity in experiment 1. Positive and negative mistunings are combined. In each panel, the parameter is $F_L$. 

Downloaded 18 Apr 2011 to 193.50.102.40. Redistribution subject to ASA license or copyright; see http://asapl.org/journals/doc/ASALIB-home/info/terms.jsp
panel, the parameter is $F_L$. The figure clearly shows that the detectability of inharmonicity monotonically worsened as $F_L$ increased from 300 to 2000 Hz. This was true for each subject, although HP did not behave exactly like LD and CS: For LD and CS, a large drop in detection performance occurred only when $F_L$ was 2000 Hz, whereas for HP, the detectability of inharmonicity became very poor as soon as $F_L$ reached 1000 Hz.

II. EXPERIMENT 2

The aim of experiment 2 was to assess the detectability of (carrier) frequency changes in the higher component of pure-tone dyads similar to those used in experiment 1. These frequency changes were detectable on the basis of pitch cues.

A. Method

On each trial, the subject heard four successive dyads with the same value of $F_L$. These four stimuli were temporally arranged in two pairs: Stimuli 1 and 2 and stimuli 3 and 4 were separated by a 300-ms pause whereas stimuli 2 and 3 were separated by 500 ms. Three of the four stimuli were identical; the odd stimulus, differing from the other three with respect to $F_H$, was the second of one of the two pairs, selected at random, and the subject had to judge which pair by pressing one of two labeled keys. All the stimuli were physically inharmonic: The two different values of $F_H$ used on any trial had a geometric mean corresponding exactly to $2F_L$. From the three identical stimuli to the odd one, $F_H$ differed by $\pm 0.15\%$, $0.3\%$, $0.5\%$, $1\%$, $2\%$, $3\%$, or $4\%$.

In all other respects (level and duration of the stimuli, selection and randomization of $F_L$ values, background noise, organization of the trials within blocks, subjects, amount of data collected, ...), experiment 2 was methodologically identical to experiment 1.

B. Results and discussion

In order to assess the stability of subjects’ behavior over time, the data obtained during the first 80 blocks of trials and the next and final 80 blocks were analyzed separately. For each of these two sets of data, each subject and each (nominal) value of $F_L$, a psychometric function similar to those shown in Fig. 3 was computed. Then, the 75 percent correct point of each function was estimated by linear interpolation.

The “thresholds” corresponding to these 75 percent correct points are symbolized by open circles in Figs. 4–6. In these figures, the upright triangles symbolize thresholds computed in exactly the same way from the data of experiment 1. The meaning of the inverted triangles will be discussed in Sec. III B.

Clearly, thresholds were lower in experiment 2 than in experiment 1. But even more importantly, $F_L$ had a different effect on subjects’ detection performances in the two experiments, as confirmed by an analysis of variance performed on the logarithms of the thresholds [subjects \* $F_L$ \* experiments: $F(3,6) = 28.0, P < 0.001$]. In experiment 1, when $F_L$ varied from 300 to 2000 Hz, thresholds monotonically increased, by an average factor of 5.8. In experiment 2, thresholds decreased for all subjects when $F_L$ went from 300 to 500 Hz, and, from 300 to 2000 Hz, they increased by an average factor of only 1.6. The average effect of $F_L$ in experiment 2 could be roughly predicted from the effect of standard frequency on the frequency discrimination of isolated and steady pure tones at a sensation level of about 25 dB (see, e.g., Nelson et al., 1983).

We conclude that for the dyads of pure tones used in both experiments, the detection of inharmonicity was not limited by imprecisions in the sensory encoding of the component tones, but by quite different and presumably more central factors.
III. EXPERIMENT 3

In experiment 1, we found better inharmonicity detection performances for negative mistunings than for positive mistunings. However, since $F_L$ was fixed within each trial, this could partly reflect response biases based on relative pitch, as discussed in Sec. I B. Experiment 3 was essentially a replication of experiment 1, but differed from it in that $F_L$ varied within each trial. The variation of $F_L$ was such that, on each trial, the stimulus with the lower pitch and the stimulus with the higher pitch had about the same probability of being positively or negatively mistuned.

A. Method

For the two stimuli presented on each trial, $F_L$ was randomly selected within two different log frequency ranges, extending respectively from 91%-96% and from 104%-109% of its nominal value. These two ranges were used at random for the first or second stimuli, but the two $F_L$ values used on a given trial were always selected in different ranges; thus, they had a minimum ratio of 104/96 = 1.083 (and a maximum ratio of 109/91 = 1.198).

In all other respects, experiment 3 was methodologically identical to experiment 1.

B. Results and discussion

Figure 7 shows the psychometric functions obtained for positive and negative mistunings, in the same format as Fig. 2. Again, negative mistunings (closed circles) were systematically more often identified as inharmonics than positive mistunings (open circles), with only one exception (subject HP for $F_c = 300$ Hz). However, it is clear that the data from experiment 3 were not identical to those from experiment 1. In the present experiment, for instance, none of the psychometric functions went significantly below the chance level, in contrast to what happened in experiment 1. This suggests that response biases based on relative pitch cues were indeed present in experiment 1. In the present experiment, however, such cues could not significantly favor the detection of negative mistunings. Therefore, the better detection of negative than positive mistunings does not seem to be artefactual.

Positive and negative mistunings were combined in order to compute "thresholds" in the same manner as for the previous data. The thresholds obtained are symbolized by inverted triangles in Figs. 4-6. For subject LD (Fig. 4), the thresholds were almost identical to those found in experiment 1. However, for $F_L = 2000$ Hz, CS and HP (Figs. 5 and 6) had better thresholds than in experiment 1. An analysis of variance performed on the logarithms of the thresholds (subjects *$F_L$ *experiments) shows that the overall performances obtained in experiments 1 and 3 did not differ significantly from each other [$F(1,2) = 2.0, p > 0.10$], but that the effect of $F_L$ was reliably different in these two experiments [$F(3,6) = 5.2, p < 0.05$].

There are two possible explanations for the improvement observed in CS and HP when $F_L = 2000$ Hz. The first one is simply that CS and HP became more sensitive to inharmonicity per se after their participation in experiments 1
and 2; in other words, these subjects would better perceive the dissonance of inharmonic dyads due to a practice effect similar in nature to those commonly found in detection or discrimination experiments. The alternative hypothesis is that the improvement is the consequence of a particular learning artefact, already described in Sec. I A 1. In experiment 2, the maximum mistuning of the stimuli corresponded to ±2% (half of ±4% since $F_n$ varied symmetrically around $2F_c$ on each trial). Thus, in this experiment, subjects were systematically presented with quasiharmonic stimuli. This may have led them, in the next experiment, to perceive the stimuli with extreme mistunings (±9%) as less familiar than the harmonic stimuli. For $F_c = 2000$ Hz, extremely mistuned stimuli may then have been identified as inharmonic not on the basis of genuine inharmonicity cues but rather from recently learned familiarity cues. For $F_c \lessapprox 2000$ Hz, such familiarity cues could not play an equally important role since genuine inharmonicity perception, as assessed in experiment 1, was possible for smaller mistunings.

IV. GENERAL DISCUSSION

For stimuli made up of two simultaneous and coherent frequency-modulated pure tones, which do not interact in the cochlea and are about 1 oct apart, we have shown the following.

1) As the carrier frequency of the lower component tone increases from about 300 to 2000 Hz, the detection of inharmonicity becomes more and more difficult (experiments 1 and 3); but although inharmonicity becomes difficult to detect as such, changes in the carrier frequency of the higher component tone do not become markedly less detectable (experiment 2).

2) Inharmonicity is more often identified as such for negative than for positive mistunings (experiments 1 and 3).

With respect to the first point, one might hypothesize that changes in the higher carrier frequency were detected on the basis of loudness cues, which are not relevant to the perception of harmonicity. Indeed, at very high frequencies, loudness cues can contribute to the frequency discrimination of steady pure tones with equal amplitudes (Henning, 1966; Coninx, 1978). However, in normal listeners, this seems to be true only for standard frequencies significantly above 4000 Hz, the (nominally) maximal carrier frequency used in the present study (Emmerich et al., 1989). Another argument against the possible role of loudness cues is that we used a background noise in all experiments; the sensation level of the tones was determined by the noise rather than by the subjects' absolute thresholds; as pointed out by Carlyon and Stubbs (1989), this minimizes the influence of the imperfect frequency response of the earphone on the sensation level of tones with different frequencies since, in any one frequency region, the earphone attenuates noise and tone alike.

Thus, it can hardly be doubted that the perceptual cues used in experiment 2 were pitch cues. Introspectively, just-noticeable changes in $F_n$ were not perceived as such, i.e., as changes in "spectral" pitch, but rather as changes in the "virtual" pitch of the whole stimulus (or more precisely as changes in its average virtual pitch since the stimuli were frequency modulated). This is consistent with the claim by Moore and Glasberg (1990, experiment 2) that virtual pitch is the predominant perceptual cue in the frequency discrimination of partials of harmonic or quasiharmonic complex tones.

From de Boer’s "phase rule" (de Boer, 1961, 1976) as well as data reported by Moore et al. (1985a, 1985b, 1986) and Hartmann (1988, p. 637), it was already clear before the present study that frequency changes producing deviations from harmonicity can be heard as pitch changes without being perceived as deviations from harmonicity as such. But the present study did not merely confirm this fact. For the dyads of pure tones used as stimuli, it showed that just-detectable inharmonicities correspond to different numbers of difference limens for pitch in different frequency registers. This suggests that the detection of inharmonicity depends on auditory processes which differ from the mechanisms of pitch discrimination.

In a recent study on the ability of human listeners to hear out a mistuned partial in an otherwise periodic complex tone with many spectral components, Hartmann et al. (1990) got certain results that resemble those obtained here in experiments 1 and 3. Their subjects progressively lost the ability to hear out mistuned partials when the frequency of the mistuned partials varied from about 2000 to 4000 Hz. This was interpreted by Hartmann et al. as the consequence of a disappearance of neural synchrony at high frequencies. However, with respect to our own data, if the decline of inharmonicity detection at high frequencies was also interpreted as the consequence of a disappearance of neural timing cues, one would be forced to conclude in addition that fine pitch discriminations (those measured in experiment 2) are not based on neural timing cues; such a conclusion may seem unrealistic in view of the present knowledge of pitch perception (see Demany, 1989, for a recent review).

In the same vein as Hartmann et al. (1990), Meddis and Hewitt (submitted) suggest that harmonic relations between components of complex tones are recognized by comparing the autocorrelation functions of the spike trains produced by these components in the auditory nerve. In the case of harmonically related components, these autocorrelation functions should have peaks for common delays; thus the detection of inharmonicity could rest on the detection of noncoincidences between autocorrelation peaks. If one assumes (Meddis, personal communication) that just-noticeable noncoincidences correspond to a fixed temporal separation between peaks, then thresholds for inharmonicity detection, when expressed in frequency percentages, should worsen when frequency increases. This is what we found in experiments 1 and 3. In experiment 2, however, it was found that thresholds for pitch discrimination were lower and did not vary in the same way with frequency; this would suggest that a different kind of neural limit was involved here.

Another way to interpret the results of experiments 1 and 3 is to consider the perception of harmonicity as a product of learning phenomena, in accordance with ideas of Terhardt (1974; see also Parncutt, 1989). Terhardt hypothesized that human listeners come to perceive two simultan-
vious pure tones 1 oct apart as a consonant stimulus because, from the beginning of their life, they hear such stimuli within the vowels of speech. Two factors might then explain the disappearance of octave harmony at high frequencies. First, the first two partials of vowels, one octave apart, are most often heard in a low register since the fundamental frequency of vowels very rarely exceeds 500 Hz. Second, although vowel spectra contain high-frequency partials among which some are 1 oct apart, these high-frequency partials are generally not resolvable from their neighbors in the auditory periphery and thus are difficult to hear out individually (Plomp, 1964; Plomp and Mimpen, 1968).

With respect to our finding that the dyads were more readily heard as inharmonic for negative than for positive mistunings, a definite explanation does not seem to be available. Hartmann et al. (1990) found that within spectrally rich quasiharmonic complex tones, the pitch of partials that are mistuned positively tends to be higher than the pitch of physically identical pure tones in isolation, while for negative mistunings pitch shifts in the opposite direction occur. These authors did not measure the pitch of nonmistuned partials. However, Terhardt (1970) found that the pitch interval formed by the two component tones of a perfectly harmonic octave complex was larger than when the same tones are heard in isolation. Thus, in the light of Hartmann et al.'s data, this pitch interval should be closer to the interval obtained for a given positive mistuning of the complex than to the interval obtained for a negative mistuning of the same absolute value. This could explain the asymmetry found here, if one admits that the internal template of the harmonic octave does not correspond to a given frequency ratio but rather to a given interval between spectral pitches. Unfortunately, in contrast to Terhardt, Peters et al. (1983) found no significant pitch shifts in the individual components of perfectly harmonic complex tones.

Let us note finally that the present study has implications with respect to the origin of human listeners' melodic octave templates. Demany and Semal (1990) observed that above 1000-1500 Hz, melodic octaves made up of successively presented pure tones can be adjusted with a much better accuracy than harmonic (i.e., simultaneous) octaves. This result put into question Terhardt's hypothesis according to which melodic and harmonic octave templates originate from one and the same learning process, taking place in infancy and based on the perceptual analysis of complex sounds such as vowels (Terhardt, 1974): If melodic octave templates were derived from experience with harmonic complex sounds, how could they be more accurate than harmonic octave templates? However, in defense of Terhardt's hypothesis, one could imagine that the inaccuracy of the high-frequency harmonic adjustments was due to a noisy frequency encoding of the two simultaneous tones, the assumed noise stemming from the tones' simultaneity. Of course, if such a noise was present from the beginning of life, it would be impossible to acquire accurate high-frequency octave templates through Terhardt's hypothetical learning process. But the results of Demany and Semal (1990) could be accounted for by assuming that the encoding noise appears after this learning process. Actually, the second experiment reported here indicated that for normal-hearing adults, the assumed high-frequency encoding noise does not exist. Thus, Terhardt's hypothesis on the origin of melodic octave templates seems to be very hard to maintain now, at least for high frequencies.

ACKNOWLEDGMENTS

This work was supported by a grant from the French Ministère de la Recherche et de la Technologie (Action Spécifique "Sciences de la Cognition") and the Conseil Régional d'Aquitaine. We thank Alain de Cheveigné and Ray Meddis for their comments on a previous version of the manuscript. Author LD is affiliated with the Centre National de la Recherche Scientifique. Author RC was supported by a Royal Society University Research Fellowship.


