Dichotic fusion of two tones one octave apart: Evidence for internal octave templates\textsuperscript{a)}

Laurent Demany
Laboratoire de Psychologie Expérimentale, Associé au C. N. R. S., Université René Descartes, 28 rue Serpente, 75006 Paris, France

Catherine Semal
Laboratoire de Psychoacoustique, U. E. R. des Sciences Sociales et Psychologiques, Université de Bordeaux II, 33405 Talence, France and Laboratoire de Psychologie Expérimentale, Associé au C. N. R. S., Université René Descartes, 28 rue Serpente, 75006 Paris, France

(Received 16 July 1987; accepted for publication 30 October 1987).

Stimuli consisting of two simultaneous and sinusoidally frequency-modulated pure tones were dichotically presented to four listeners. The two component tones of each stimulus were approximately an octave apart. They were both modulated at 2 Hz, and the frequency swing resulting from each modulation corresponded to one tenth of the carrier frequency. The listeners’ task was to detect phase differences between the modulation waveforms of the two simultaneous tones: With an adaptive 2IFC procedure, just-noticeable values of $\Phi$, the phase angle of the modulation waveforms, were measured as a function of the interval formed by the carrier frequencies (one octave, i.e., 1200 cents, $\pm 0, 25, 50,$ or 100 cents). When the carrier frequencies were not too high, just-noticeable values of $\Phi$ often varied nonmonotonically with the interval, showing a minimum at or near 1200 cents. An additional experiment indicated that most, if not all, of these octave effects were not due to some form of beat detection. As a whole, the results reported here provide evidence for the existence of internal octave templates. Such templates might play an important role in the perceptual segregation of simultaneous harmonic signals, as well as in pitch perception.

PACS numbers: 43.66.Hg, 43.66.Rq, 43.66.Lj [NFV]

INTRODUCTION

Any complex and periodic sound signal, with period $p$, can be described as a sum of sinusoidal tones with frequencies equal to $n_k/p$, where $n_k$ is an integer. The signal is said to be “harmonic” when $n_k$ is always equal to $n_{k-1} + 1$, or in other words, when the frequency ratio of two adjacent partials is always the ratio of two successive integers (2/1, 3/2, etc.). This condition is not satisfied for “inharmonic” signals. Generally speaking, harmonic signals are not perceived in the same way as inharmonic signals. Harmonic signals tend to evoke a single, clearly defined low-pitch sensation; their partials fuse into a coherent sound image. By contrast, inharmonic signals do not, in general, evoke a single pitch, and their partials tend to be more easily heard individually (Bregman and Doehring, 1984; Grandori, 1984; Martens, 1984; McAdams, 1982; McAdams et al., 1985, 1986).

This perceptual difference between harmonic and inharmonic signals can be accounted for in the framework of the “pattern recognition” theories of low-pitch perception, especially Terhardt’s and Goldstein’s theories (Terhardt, 1974; Terhardt et al., 1982; Goldstein, 1973; Duifhuis et al., 1982; Scheffers, 1983). Both of these theories assume the existence, in the central auditory system, of “internal templates” representing specific intervals between partials. The intervals corresponding to these hypothesized templates are the frequency ratios of the partials of harmonic signals, which implies that human listeners would possess an internal octave, an internal fifth (or twelfth), etc. It can be argued that the sine tones contained in steady harmonic signals fuse into one sound image because the intervals they form match internally defined templates (see McAdams, 1984).

But do such internal templates really exist? This remains, in fact, to be demonstrated objectively. Psychophysical evidence for the existence of internal octave templates was looked for in the present research.

Evidence for the existence of internal harmonic templates can be sought in the following way. Assume that a human listener is required to detect small changes in the interval formed by two simultaneous sine tones, that is, to discriminate between slightly different frequency ratios. Assume, in addition, that the listener possesses, for instance, an internal octave template but no internal template corresponding to a seventh or a ninth. It might then be expected that the listener’s performance in the task will be better if the standard frequency ratio corresponds to an octave than if it corresponds to a seventh or a ninth.

The discriminability of frequency ratios has recently been studied by Viemeister and Fantini (1987). Their data indeed show that deviations from an octave interval are easier to detect than deviations from a somewhat smaller or larger interval. In their experiments, however, the two tones forming each stimulus were presented monaurally and at a rather high level (67.5 dB SPL). In such conditions, the

\textsuperscript{a)}Part of this article was presented at the 112th Meeting of the Acoustical Society of America, Anaheim, CA [J. Acoust. Soc. Am. Suppl. 1, 80, S91 (1986)].
observed octave effect might originate from peripheral interactions between the tones and not from the existence of internal octave templates. When two pure tones of moderate or high level and about one octave apart are simultaneously presented to the same human ear, some cochlear fibers of the listener certainly respond to both of these tones [see the data obtained in the squirrel monkey by Rose et al. (1971), and in the cat by Greenberg (submitted for publication)]. If the two tones are separated by exactly one octave, the effective waveform driving these fibers will have a very short period (the period of the lower frequency tone). But the waveform period will be much longer if the two tones form a slightly different interval, and this may be signaled perceptually by a beat sensation (cf. Plomp, 1976, Chap. 3).

In order to provide evidence for the existence of internal harmonic templates by means of discrimination experiments, it is, therefore, necessary to prevent peripheral interactions between partials. In the present research, which attempted to demonstrate the existence of octave templates, peripheral interactions were prevented by presenting partials dichotically and at low level.

I. EXPERIMENT 1

A. Method

1. Stimuli and rationale

In this first, main experiment, each stimulus consisted of two simultaneous, dichotically presented, and slowly frequency-modulated pure tones. The two tones had different carrier frequencies; the lower carrier frequency will be labeled \( F_L \) and the higher one \( F_H \). For each tone, the instantaneous frequency varied sinusoidally at a rate of 2 Hz, and the total frequency swing was 10% of the carrier frequency. Two successive stimuli were presented on each trial. For one of the two stimuli (stimulus a; see Fig. 1, upper part), the modulation waveforms of the two tones were in phase; thus, while each tone varied in frequency, the frequency ratio of the two tones had a steady value, equal to the ratio of their carrier frequencies (\( F_H / F_L \)). The two tones forming the other stimulus (stimulus b; Fig. 1, lower part) had the same carrier frequencies, \( F_L \) and \( F_H \), but their modulation waveforms were not in phase; thus their frequency ratio was not steady, but varied in a quasisinusoidal way around a central value corresponding to \( F_H / F_L \). On each trial, the listener had to decide whether stimulus b had been presented in the first or second position. Note that for every stimulus presentation, the initial phases of the two modulation functions were random variables. This prevented the subjects from using monaural pitch cues to perform the task.

Within trial blocks, the phase difference \( \Phi \) between the two modulations contained in stimulus b was varied, and just-noticeable values of \( \Phi \) were measured. The logic of this measure lies in the fact that, in stimulus b, the amplitude of the frequency ratio fluctuation (expressed in cents) increases quasilinearly with \( \Phi \) when \( \Phi \) goes from 0° to about 70°. Thus measuring a just-noticeable value of \( \Phi \) was, in fact, equivalent to measuring a just-noticeable frequency ratio fluctuation. Figure 2 shows the exact relation between \( \Phi \) and the frequency ratio fluctuation. It must be emphasized that this relation does not depend on the standard frequency ratio, i.e., \( F_H / F_L \).

Our main independent variable was, of course, \( F_H / F_L \). This ratio took seven possible values: one octave (1200 cents) ± 0, 25, 50, and 100 cents. In addition, the absolute values of \( F_L \) and \( F_H \) were varied over a wide range. Generally, the tone with the lower carrier frequency (\( F_L \)) was heard through the right ear, but we looked for possible ear asymmetries by reversing, in some experimental conditions, the position of the earphones. All tones but one were presented at 45 dB SPL. The exception was the tone with the lowest of the carrier frequencies used, 180 Hz; it was presented at 50 dB SPL.

Each stimulus had a steady-state portion of 3 s and 200-ms rise/fall times, which were shaped with a raised-cosine function. The two stimuli presented in every trial were separated by a 200-ms silent interval. Stimuli were heard in an IAC soundproof room through TDH 49 earphones.
In each experimental session, $F_L$ was fixed and seven thresholds were measured, one for each possible value of $F_H/F_L$. These seven thresholds were measured in a random order, varying from session to session and unknown to the subject throughout each session. For a given value of $F_L$ (and a given position of the earphones), each subject was tested for at least six successive sessions: a variable number of training sessions plus six data collection sessions. All sessions were run on different days. For each subject, at least three values of $F_L$ were used. They were extracted from the following set of frequencies: 180, 270, 400, 600, 900, and 1350 Hz.

3. Subjects

Four subjects, aged 22 to 32, participated in the experiment. All had a normal audiogram (thresholds better than 15 dB HL) from 250 to 4000 Hz. Subjects 1 and 4, the authors, had participated as subjects in previous psychoacoustic experiments; this was not the case for subjects 2 and 3. Subjects 1 and 3 were amateur musicians, but subjects 2 and 4 were relatively naive musically.

B. Results

The results are displayed in Figs. 3–6. Each panel of these figures shows the mean thresholds obtained in a given
subject for a given value of $F_L$ and a given position of the earphones.

We wanted to know, mainly, if thresholds would be a nonmonotonic function of $F_H/F_L$, with a minimum at or near the octave (1200 cents). To answer this question, the data presented in each panel were submitted to an analysis of variance where $F_H/F_L$, the single factor, was considered as a "treatment" factor (cf. Winer, 1971, Chap. 4). The use of analyses of variance was justified since there was no evidence that the measures were serially dependent (Shine and Bow-er, 1971). In each analysis, an overall $F$ statistic, testing the existence of differences between the mean thresholds, was first computed. If the computed $F$ allowed for a rejection of the null hypothesis [$F(6,30) > 2.87; p < 0.025$], the existence of a significant quadratic trend was then tested using the appropriate contrast between means. In addition, the existence of a significant linear trend was also tested by means of the appropriate contrast.

Table I presents the results of these analyses. Clearly, the results depend on both $F_L$ and the subject. For each subject, the overall $F$ test was significant for at least one value of $F_L$ and at least one reliable quadratic trend was
### TABLE I. F statistics computed from the data obtained in experiment 1. Earphone position a: lower frequency tone at the right ear; earphone position b: lower frequency tone at the left ear. ***: p < 0.001; **: p < 0.01; *: p < 0.025; ns: p > 0.025.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Earphone position</th>
<th>Degrees of freedom</th>
<th>180</th>
<th>270</th>
<th>400</th>
<th>600</th>
<th>900</th>
<th>1350</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>Overall</td>
<td>6.30</td>
<td>6.1</td>
<td>22.9</td>
<td>33.5</td>
<td>1.3</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quadratic trend</td>
<td>1,5</td>
<td>5.8</td>
<td>111.5</td>
<td>57.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Linear trend</td>
<td>1,5</td>
<td>14.2</td>
<td>0.1</td>
<td>87.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Overall</td>
<td>6.30</td>
<td>11.3</td>
<td></td>
<td>0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quadratic trend</td>
<td>1,5</td>
<td>24.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Linear trend</td>
<td>1,5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 2       | a                | Overall           | 6.30| 7.5 | 9.6 | 7.1 |     |      |
|         |                  | Quadratic trend   | 1,5 | 49.8| 37.7| 9.8 |     |      |
|         |                  | Linear trend      | 1,5 | 0.0 | 1.0 | 133.0|     |      |
|         | b                | Overall           | 6.30| 2.0 |     | 11.9| 2.5  |      |
|         |                  | Quadratic trend   | 1,5 |     |     | 122.4|     |      |
|         |                  | Linear trend      | 1,5 |     |     | 11.0 |     |      |

| 3       | a                | Overall           | 6.30| 20.5| 24.8| 9.3 | 5.6 | 2.4  |
|         |                  | Quadratic trend   | 1,5 | 38.4| 150.0| 13.9| 32.6|      |
|         |                  | Linear trend      | 1,5 | 1.3 | 1.0 | 0.5 | 0.2  |      |
|         | b                | Overall           | 6.30| 33.9|     | 1.4 |     |      |
|         |                  | Quadratic trend   | 1,5 | 234.8|     |     |      |      |
|         |                  | Linear trend      | 1,5 |     |     | 27.2 |     |      |

found. On the whole, more than half of the analyses demonstrated the existence of a significant quadratic trend. However, the overall F test was not always significant, and the subjects often did not behave in one and the same way at the same value of $F_L$. Note that significant linear trends were found in five cases; these unexpected trends, each reflecting a threshold decrease when $F_H/F_L$ increases, are difficult to interpret.

With respect to the significance of the overall F test and the presence or absence of a reliable quadratic trend, the position of the earphones appeared to have no influence. Figures 3–6 and Table I suggest that, for each subject, quadratic variations of the threshold with $F_H/F_L$ occur if and only if $F_L$ and/or $F_H$ are located within some limited frequency domain. The probable limits of the individual domains are shown in Fig. 7. A lower limit could not be estimated for subjects 2 and 4, but an upper limit seems to exist for each subject. When expressed as a function of $2F_L$, this upper limit seems to vary from about 1000 Hz in subject 1 to about 2000 Hz in subject 4.

![FIG. 7. Frequency domains within which statistically significant octave effects (i.e., quadratic threshold variations) were found in experiment 1. The domains are represented by dotted lines. Horizontal dashes indicate the existence of an octave effect and vertical dashes indicate the absence of an octave effect. For subjects 1 and 4, the upper dashes refer to a right-ear presentation of the lower frequency tone and the lower dashes refer to a left-ear presentation of this tone.](image-url)
The function obtained by averaging the 20 functions displayed in Figs. 3–6 has a minimum for $F_h/F_L = 1200$ cents and is nearly symmetrical. However, when thresholds varied nonmonotonically with $F_h/F_L$, the minimum threshold was sometimes obtained not for $F_h/F_L = 1200$, but rather 25 or 50 cents lower or higher. Consider, in this respect, the circled data points in Figs. 3 and 6. Each of these three points represents a mean threshold that is significantly smaller ($p < 0.025$ on a Student t test) than the mean threshold obtained during the same sessions—for $F_h/F_L = 1200$ cents. Admittedly, it is difficult to draw firm conclusions from post-hoc statistical comparisons. Nevertheless, the small “mistunings” just noted might be meaningful: Since each stimulus was dichotic, they might originate from binaural diplacusis and reflect small deviations of the subjective octave from the physical octave (see, e.g., van den Brink, 1974).

C. Discussion

Experiment 1 was performed to provide evidence for the existence of internal octave templates, and the results indeed suggest that such internal templates exist. However, another possible interpretation of the octave effects found must be considered immediately. Although the two tones forming each stimulus had low SPLs and were presented dichotically, it might still be imagined that the octave effects stem from some form of beat detection. Assume, for instance, that in the experimental situation, significant aural harmonics were produced at the ear receiving the lower frequency tone (carrier frequency $F_L$). For $F_h/F_L = 1200$ cents (one octave), one of these potential aural harmonics has the same instantaneous frequency as the contralateral tone as long as $\Phi = 0^\circ$, and this is no longer true when $\Phi \neq 0^\circ$. Thus ordinary lateralization cues can, in principle, be used to detect small phase differences between the modulation functions of the two tones.

With regard to this hypothesis, it should be noted first that the subjects, in fact, never reported hearing beats or “rotating” sound images when the frequency ratio of the two tones was fluctuating, i.e., when $\Phi$ differed from $0^\circ$. For all the conditions in which octave effects on performance were observed, these effects seemed to be, introspectively, the direct consequence of spectral fusion phenomena and not the by-product of some form of beat detection. Assume, for instance, that in the experimental situation, significant aural harmonics are produced at the ear receiving the lower frequency tone (carrier frequency $F_L$). For $F_h/F_L = 1200$ cents (one octave), one of these potential aural harmonics has the same instantaneous frequency as the contralateral tone as long as $\Phi = 0^\circ$, and this is no longer true when $\Phi \neq 0^\circ$. Thus ordinary lateralization cues can, in principle, be used to detect small phase differences between the modulation functions of the two tones.

In each trial, they voted for the stimulus in which the two tones appeared to fuse less or, in other words, were more difficult to hear as a single sound entity. Even when $F_h/F_L$ was equal to 1100 or 1300 cents, the two tones fused better for $\Phi = 0^\circ$ than for $\Phi \neq 0^\circ$; fusion was promoted when the two modulations were coherent, as could be anticipated from previous research (Bregman and Doehring, 1984; McAdams, 1984). But when $F_h/F_L$ was equal or very close to 1200 cents, $\Phi$ had a more pronounced effect on perceptual fusion because the two tones fused particularly well for $\Phi = 0^\circ$.

In each session, $f$ was fixed and eight blocks of 50 trials “mistunings” pointed out in the preceding section: If aural harmonics were really responsible for the observed octave effects, then thresholds should have always been minimum for $F_h/F_L = 1200$ cents. Actually, the tips of the U-shaped functions obtained were sometimes located a bit higher or lower. This seems to rule out not only the aural harmonics hypothesis, but also the possible intervention of artifactual monaural interactions due to interaural cross talk.

It may be considered, however, that the small mistunings just mentioned, as well as phenomenological arguments, do not rule out convincingly enough uninteresting explanations for the octave effects found in experiment 1, i.e., the possibility of some form of beat detection. In order to provide more support for the “interesting” explanation—the octave template hypothesis—a second experiment was performed. Its logic was as follows. Assume that, in experiment 1, a given subject detects some form of beat when $F_L$ has a certain value ($f$, $F_h = 2f$, and $\Phi \neq 0^\circ$). The detection of the beat implies a sensitivity to the relation between the temporal fine structures of the two monaural waveforms: No beat could be detected if only spectral information (i.e., fluctuating pitches) was extracted from the waveform of each tone. Assume, therefore, that the two modulated tones are replaced by two steady pure tones with frequencies respectively corresponding to $f$ ($= F_h$) and $2f$. The beat detection hypothesis implies that the subject will be able to detect changes in the interaural phase relation of these two steady tones. The aim of experiment 2 was to determine if such changes were indeed detectable.

II. EXPERIMENT 2

A. Method

This experiment was conducted on the four subjects already used in experiment 1.

Each stimulus consisted of two simultaneous and dichotically presented pure tones, with frequencies $f$ (at the right ear) and $2f$ (at the left ear). For each subject, the value taken by $f$ was always a value of $F_L$ for which a quadratic threshold variation (i.e., an octave effect) had been obtained in experiment 1. The two tones forming each stimulus had the same SPL as the corresponding tones in experiment 1. They were generated at a sampling rate of 43.8 kHz, with the same equipment and the same precision as in experiment 1. Each stimulus had a steady-state portion of 1 s and 200-ms rise/fall times, shaped with a raised-cosine function.

In each trial, four successive stimuli were presented. They were arranged in two successive pairs: 200-ms silent intervals separated the two stimuli of each pair, and a 600-ms silent interval separated the two pairs. Three of the four stimuli were identical; their two component tones were in sine phase at the input of the earphones. The remaining stimulus had the same component tones, but in a different phase relation: In some trial blocks, the relative phase of the higher frequency tone was shifted by $+90^\circ$; in other trial blocks, the phase shift was $+180^\circ$. The “different” stimulus was always the second of a pair, and the subject had to decide, on each trial, if it occurred in the first or second pair. Feedback was provided exactly as in experiment 1.

In each session, $f$ was fixed and eight blocks of 50 trials
were run (four blocks with each of the two phase shifts). Two or three successive sessions were run, on different days, for a given value of \( f \).

### B. Results and discussion

Table II shows the percentages of correct responses obtained in each subject for each value of \( f \) and each phase shift. Random responses would, of course, yield 50%-correct performance. It can be seen that performance is at the chance level in subjects 2 and 3, but often exceeds the chance level in subjects 1 and 4. Even in subjects 1 and 4, however, performance is generally quite poor: Except for subject 4 when \( f = 180 \text{ Hz} \), the percentages of correct responses never exceed 65%. Subjects 1 and 4 felt that the phase shifts were, at times, detectable on the basis of an extremely subtle change in the "timbre" of the standard stimulus; but they also felt that phase shifts were completely impossible to detect on the basis of lateralization criteria.

The authors of two previous studies (Craig and Jeffress, 1962; Ayres and Clack, 1984) reported that at low sensation levels, human listeners cannot reliably discriminate between dichotic octave complexes that differ only by the phase relationship of their two component tones. As a whole, the results of experiment 2 confirm these findings. Conflicting claims have been made by Thurlow and Bernstein (1957) and Tobias (1964): According to these authors, a beat sensation can be evoked in at least some subjects by a low-level dichotic stimulus that consists of two steady pure tones forming a mistuned octave (e.g., \( 2f + 1/f \)). But it might be that the tones used in these two studies suffered, in fact, from harmonic distortion. In an informal experiment, we produced dichotic mistuned octaves with two steady pure tones at the same SPLs as in experiments 1 and 2. No beat was audible to us, although we were, in experiment 2, the two subjects performing above the chance level.

We thus conclude from experiment 2 that most—if not all—of the octave effects found in experiment 1 do not stem from some form of beat detection.

### III. GENERAL DISCUSSION

The present findings tally very well with the idea that octave templates exist in the central auditory system of man. Strong support for that idea was provided in so far as (i) our data consist of discrimination performances, and are thus as "objective" as psychophysical data can be, and (ii) results supporting the octave template hypothesis were obtained from each of the four subjects tested in the main experiment (experiment 1). In connection with the latter point, it is worthy to note that in experiment 1, the most pronounced octave effects were found in the least "musical" subject (subject 4); this can be taken as evidence that the octave effects found are truly sensory effects and should not be attributed to a musical acculturation process.

The existence of internal templates for harmonic frequency ratios is a central assumption of Goldstein's and Terhardt's "pattern recognition" theories of pitch perception. In so far as our results directly support this central assumption, they indirectly support both of the theories. However, the present study was concerned only with the existence of octave templates, while Goldstein's and Terhardt's theories assume the existence of other harmonic templates. Among all the harmonic frequency ratios, the octave may be "special" (Bachem, 1950; Demany and Armand, 1984). We are presently determining if, with the method used in experiment 1, evidence for the existence of internal templates corresponding to frequency ratios such as 3/1, 3/2, or 4/1 can be provided.

Our results indirectly support pattern recognition models of pitch perception, but it must be acknowledged that they are also compatible with a pitch theory that does not assume that the extraction of pitch from complex tones is based on a pattern recognition process (Moore, 1982, Sec. 4.4.; van Noorden, 1982). Following Moore and van Noorden, it may be speculated that the central auditory system detects harmonic relationships between two simultaneous pure tones by comparing neural timing informations conveyed by remote cochlear fibers. The neural spikes produced in cochlear fibers by a pure tone with period \( p \) and a pure tone with period \( 2p \) produce neural spikes that are separated by common time intervals. The central auditory system might detect the harmonic relationship of the two tones on this basis. Since Moore (1982) and van Noorden (1982) admit that the harmonic relationship of two simultaneous pure tones can be recognized in the absence of peripheral interactions between these tones, our results do not conflict with their speculations.

In Sec. I C, we stated that throughout experiment 1, the response criterion used by the subjects was the amount of fusion of the two tones forming each stimulus. It is interesting to consider here an alternative hypothesis based on the well-known results of Houtsma and Goldstein (1972).
Houtsma and Goldstein demonstrated that two simultaneous and dichotically presented pure tones that are successive harmonics of some missing fundamental are able to evoke a "central" pitch corresponding to the missing fundamental. With regard to experiment 1, it might thus be imagined that: (i) the two tones forming the stimuli always fused, and the amount of fusion did not depend on the stimulus; (ii) each stimulus evoked a fluctuating central pitch, determined at any moment by the instantaneous frequencies of the two tones; and (iii) on each trial, subjects based their response on the magnitude of the fluctuation of central pitch for each of the two stimuli. We applied Goldstein's theory (Goldstein, 1973) to simulate the fluctuation of central pitch for various values of $F_H/F_L$ and $\beta$: For two instantaneous frequencies $f_L$ and $f_H$ ($f_H < f_L$), the central pitch was considered to correspond to $(f_L + 0.5f_H)/2$ (Ref. 6). Two conclusions were drawn from the simulation. First, for each value of $F_H/F_L$, the amplitude of the fluctuation of central pitch (expressed as a frequency ratio) decreases when $\Phi$ deviates from $0^\circ$; second, the rate of this decrease is a monotonic function of $F_H/F_L$, and is largest for $F_H/F_L = 1100$ cents. Thus the octave effects found in experiment 1 were not predictable from the results of the simulation; these results predicted that thresholds for $\Phi$ would always decrease regularly with $F_H/F_L$. We believe that the simulation failed to predict octave effects because it was based on the assumption that the fusion of the two tones making the stimuli did not depend on $F_H/F_L$ or $\Phi$. In fact, the fusion of the two tones, and thus the saliency of central pitch, depended on these two variables.

Finally, one must ask why, in experiment 1, octave effects appeared to occur only for a limited range of $F_L$ values, varying from subject to subject. There may be some link between this phenomenon and the so-called "spectral dominance" phenomenon in pitch perception (Plomp, 1976, Chap. 7; Moore et al., 1983). However, the frequency limits of the observed octave effects may be principally due to the use of dichotic, and thus "abnormal," stimuli (see Hall and Soderquist, 1975; Raatgever and Bakkm, 1986). One of our projects is to replicate experiment 1 with monaural rather than dichotic stimuli. In experiment 1, the two tones forming the stimuli were presented dichotically in order to prevent peripheral interactions. But peripheral interactions can also be prevented in a monaural listening situation by adding to the tones a narrow-band noise spectrally centered halfway between their frequencies. It will be interesting to determine whether, in this situation, octave effects occur for larger ranges of $F_L$ values than in a dichotic listening situation.

\[ f'_L \text{ (instantaneous frequency of the lower frequency tone) and } f'_H \text{ (instantaneous frequency of the higher frequency tone) varied with time according to} \]
\[ f'_L = f_L [1 + (0.05 \sin 4\pi t)], \quad (1) \]
\[ f'_H = f_H [1 + (0.05 \sin (4\pi t + \Phi))], \quad (2) \]
\[ \text{where } \Phi = 0 \text{ or } \Phi \neq 0 \text{, as } \Phi \text{ deviates from } 0^\circ. \]
\[ \text{and the amount of fusion did not depend on the stimulus;} \]
\[ \text{since the amplitudes of the frequency variations were small, we had} \]
\[ f'_L/f'_H = \frac{F_H/F_L}{1 + (1 + 0.05 \sin (4\pi t + \Phi))} \cdot \frac{1}{(1 + 0.05 \sin (4\pi t))}. \quad (3) \]
\[ \text{This can be rewritten as} \]
\[ f'_L/f'_H = \frac{F_H/F_L}{1 + (0.1 \sin \frac{\pi}{2}) \cos \left(4\pi t + \frac{\pi}{2}\right)}. \quad (4) \]

\[ \text{Thus } f'_H/f'_L \text{ varied in a quasiphasoidal way between } (F_H/F_L) \times \left[1 - (0.1 \sin \pi/2)\right] \text{ and } (F_H/F_L) \times \left[1 + 0.1 \sin \pi/2\right]. \]

\[ \text{If the initial phase of the modulation of one of the two tones had been fixed, subjects could have performed the task by detecting changes in the initial pitch of the other tone. Randomizing the initial phases prevented this possibility. For clarity, the randomization of the initial phases was ignored in the equations of footnote 1.} \]

\[ \text{However, the maximum value that } \Phi \text{ could take was } 180^\circ; \text{ thus } \Phi \text{ was set to } 180^\circ \text{ when, according to the variation rule, it should have exceeded } 180^\circ. \text{ In fact, } \Phi \text{ very rarely reached } 180^\circ \text{ during the experiment.} \]

\[ \text{The contrast consisted of the following weighting coefficients: 10 for 1100 and } 1300 \text{ cents, } -2 \text{ for } 1150 \text{ and } 1250 \text{ cents, } -5 \text{ for } 1175 \text{ and } 1225 \text{ cents, and } -10 \text{ for } 1200 \text{ cents; this was the appropriate contrasts to test the hypothesis that the variation of thresholds with } F_H/F_L \text{ was a quadratic function with an extremum for } F_H/F_L = 1200 \text{ cents. See Winer (1971) for explanations on the rationale of statistical tests for trend.} \]

\[ \text{Subject 1 reported that when the component tones of a stimulus fused very well, these tones were sometimes heard as coming both from the same place within the head. But according to the subject's report, this phenomenon was generally produced by each of the two stimuli presented in a given trial. Thus the phenomenon could not usefully serve as a response cue. Another subject (subject 4) reported that when the tones fused very well, the higher frequency tone sometimes disappeared during the stimulus presentation. This phenomenon, which resembles the illusion described by Deutsch (1974), was exploited by the subject because she felt that it occurred mostly for } \Phi = 0^\circ \text{ and could serve as an efficient cue.} \]

\[ \text{The central pitch was computed from Eq. (12) of Goldstein's article. According to that equation, the central pitch corresponds to} \]
\[ \frac{h x_h^2}{\delta_x^2} + (h + 1) x_h^2 \delta_x^2 \]
\[ \delta_x^2 \delta_y^2 + (h + 1) \delta_x^2 \delta_y^2 \]
\[ \text{where } h \text{ represents the estimated harmonic rank of the tone with frequency } f_x, x_h \text{ and } x_y \text{ are the two aurally measured instantaneous frequencies, and } \delta_x \text{ and } \delta_y \text{ can be considered as error estimates for the aural measurement of } f_x \text{ and } f_y. \text{ For realistic values of } x_h \text{ and } x_y, i.e., values very close to } f_x \text{ and } f_y, \text{ the pitch processor will always set } \delta = 0. \text{ Then, if one assumes that } \delta_x = 2 \delta_y, \text{ the central pitch can be numerically computed as } (f_x + 0.5f_y)/2. \text{ It is somewhat arbitrary to assume that } \delta_x = 2 \delta_y. \text{ However, the qualitative conclusions we draw from the simulation do not depend on this particular assumption: For any value of } \delta_x/\delta_y, \text{ the same conclusions are reached. It should be noted that Goldstein's (1973) pitch processor always interprets the spectral components of a complex tone as successive harmonics. Gerson and Goldstein (1977) modified Goldstein's original model, and the pitch processor that they propose can also consider spectral components as non-successive harmonics. But in the present case, using Goldstein's original model to compute central pitch was justified because the subjects never heard central pitches corresponding to frequencies distant from } f_x, \text{ which implies that } f_x \text{ and } f_y \text{ were never interpreted as non-successive harmonics of some missing fundamental.} \]

\[ \text{ACKNOWLEDGMENTS} \]

We wish to thank Ervin R. Hafter, William M. Hartmann, Adrian J. M. Houtsma, Stephen E. McAdams, Robert W. Peters, and Bennett K. Smith for their comments on a previous version of this article. We are also grateful to Dominique Lépine for his statistical advice and to Hélène Carlus, who was subject 3, for her generous cooperation.

\[ \text{Bachem, A. (1950).} " \text{Tone height and tone chroma as two different pitch qualities,}" Acta Psychol. 7, 80-88. \]


