Pitch discrimination of harmonic complex signals: Residue pitch or multiple component discriminations?

Andrew Faulkner
Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, England

(Received 19 February 1985; accepted for publication 26 August 1985)

Two models for pitch discrimination of harmonic complex sounds are discussed, a multiple-band probability summation model using comparisons among component frequencies, and a model in which residue pitches are compared. The second model is based on Goldstein's optimum-processor pitch theory [J. Acoust. Soc. Am. **54**, 1496–1516 (1973)], and is distinguished from the multiple-band model by an internal noise process. Pitch difference limens from 2I2AFC tasks show that when the test signals comprise corresponding harmonics, relative pitch difference limens are less than the smaller relative difference limens for the component frequencies, which is consistent with the multiple-band model. The absence of corresponding harmonics significantly reduces relative pitch discriminability; this effect supports the model on Goldstein's theory. It appears that residue pitch comparisons are not used for pitch discrimination between sounds with corresponding components; rather, comparisons based on residue pitch are only employed where there are no common resolved components in the signals to be discriminated.

PACS numbers: 43.66.Hg, 43.66.Fe, 43.66.Ba

INTRODUCTION

Human observers' sensitivity to changes in the pitch of complex signals is of direct relevance to the perception of both speech and music, but, with some exceptions (Flanagan and Saslow, 1958; Schodder and David, 1960; Ritsma, 1963; Henning and Grosberg, 1968; Walliser, 1969; Cudahy, 1975; Fastl and Weinberger, 1981; Horst *et al.*, 1984; Moore *et al.*, 1984; Scheffers, 1984; Spiegel and Watson, 1984), studies of pitch discrimination have investigated sinusoidal stimuli (e.g., Shower and Biddulph, 1931; Turnbull, 1944; Harris, 1952, 1966; Liang and Chistovich, 1961; Henning, 1966, 1967, 1970; Moore, 1973; Weir *et al.*, 1977a). Pitch discrimination can be related directly to auditory frequency analysis (Helmholtz, 1863), and that theoretical framework suggests that the factors which affect discrimination of simple signals will affect discrimination of complex signals in the same way. In some cases, however, it is possible that additional processes, such as those involved in the perception of "residue" or "virtual" pitch, may play a part in discrimination. For example, the "optimum-processor" theory (Goldstein, 1973; Srulovicz and Goldstein, 1983) proposes that the precision of residue pitch is not limited by the discriminability of the component frequencies, but rather by noisy frequency-specific channels conveying estimates of the frequencies of components to a central pitch processor. We now briefly review two models for the discrimination of the pitch of complex signals.

I. MODELS OF DISCRIMINATION OF COMPLEX SIGNALS

A. A multiple-band model

The simplest plausible model of pitch discrimination for complex signals is a multiple-band probability summation process in which the discriminability of the pitch of a complex signal is derived directly from the discriminability of the components. From this model, which is elaborated in Appendix A, the relative DL of a complex signal $\delta(f_0)/f_0$ is given by

$$\left(\frac{f_0}{\delta(f_0)}\right)^2 = \sum_{i=1}^{N} \left(\frac{f_i}{\delta(f_i)}\right)^2, \tag{1}$$

where $\delta(f_i)/f_i$ is the relative DL of the ith component of the signal.

B. A residue pitch model

Modern pitch theories (Terhardt, 1970, 1974; Goldstein, 1973; Srulovicz and Goldstein, 1983; Wightman, 1973) lead to the prediction that the discriminability of the residue pitch will be based indirectly upon component frequency discriminability through the pitch extraction process. As de Boer (1977) has argued, these pitch theories have many similarities and should lead to similar predictions. Goldstein (1973) has derived the following prediction for the precision of residue pitch from his optimum-processor model:

$$\left(\frac{f_0}{\sigma(f_0)}\right)^2 = \sum_{i=1}^{N} \left(\frac{f_i}{\sigma(f_i)}\right)^2, \tag{2}$$

where $\sigma(f_0)/f_0$ is the relative standard deviation of the residue pitch estimate in the region of the fundamental frequency, and $\sigma(f_i)/f_i$ is the relative standard deviation of an internal estimate of the frequency of the ith component [Goldstein's Eq. (15)]. A computer simulation of Wightman's (1973) model makes a similar prediction to Eq. (2) (Faulkner, 1982). If $\sigma(f_i)/f_i$ is equivalent to $\delta(f_i)/f_i$ then Eq. (2) is equivalent to Eq. (1), the prediction of the multiple-band model.

Estimates of the model parameter $\sigma(f_i)/f_i$ from error rates in melody and musical interval identification are 0.010

© 1985 Acoustical Society of America

or more between 1 and 10 kHz (Goldstein, 1973; Goldstein *et al.*, 1978; Houtsma, 1979); these estimates are notably larger than the pure-tone relative DL, which may be smaller than 0.002 between 1 and 2 kHz (Shower and Biddulph, 1931; Henning, 1967, 1970; Moore, 1973; Weir *et al.*, 1977a). The discrepancy could be due to some unknown source of error in pitch identification, but Goldstein assumes it to be due to an internal component of $\sigma(f_i)/f_i$. If this internal component is independent of the frequency DL, $\sigma(f_i)/f_i$ may be rewritten in terms of $\delta^2(f_i)/f_i$ and an additional internal noise $\sigma^2(\text{int}_i)$ as

$$\sigma^2(f_i)/f_i = \delta^2(f_i)/f_i + \sigma^2(\text{int}_i). \qquad (3)$$

II. PREVIOUS STUDIES OF DISCRIMINATION OF COMPLEX SIGNALS

Previous studies indicate that the relative discriminability of the pitch of a complex signal may be considerably better than the discriminability of the fundamental frequency component alone, and comparable to the discriminability of the best discriminated components of the complex (Flanagan and Saslow, 1958; Henning and Grosberg, 1968; Walliser, 1969; Moore *et al.*, 1984; Spiegel and Watson, 1984).

It is unclear whether such discriminations depend upon the perception of residue pitch, but an experiment reported by Ritsma (1963) suggests that they do not. Ritsma found the relative DL for the absent fundamental frequency of a harmonic signal to be larger than that for an isolated pure tone in the region of the complex signal's components. In Ritsma's study, two signals without common components were compared, while in the other studies cited here, the two signals comprised identical series of harmonics. It is thus possible that residue pitch differences were used only in Ritsma's experiment, while in the other studies, performance was based on component frequency differences.

If the optimum-processor model's assumption of an internal noise is correct, the relative DL for residue pitch may well be, as Ritsma found, greater than the relative DLs of the components that are combined to produce the residue pitch. Furthermore, the DL for residue pitch will be unaffected by external noise until that noise is similar in power to any internal noise. The experiments described below test these predictions. In one task, the conditions allow observers to make multiple component frequency comparisons, and, if they do so, their performance should be as predicted by the multiple-band model. In a second task, component frequency comparisons do not provide a useful cue, forcing the use of residue pitch differences.

III. INTRODUCTION TO THE EXPERIMENTS

In order to observe any correlation of performance in the discrimination of simple and complex tones, two parameters known to affect the discriminability of the frequency of simple tones were varied: level (Shower and Biddulph, 1931; Harris, 1952; Weir *et al.*, 1977a) and signal-to-noise ratio (Harris, 1966; Henning, 1967). It is important to note that the components of a complex signal may mask each other. Thus the DL for a single sinusoid may not be a good estimate of the discriminability of a component of a complex signal at that same frequency; in order to avoid this problem, estimates of frequency DLs for single components are obtained with the other components of the complex signal also present as continuous tones of constant frequency. Moore *et al.* (1984) used a similar procedure, but presented the other components as pulsed tones of the same duration as the test component.

A 2-interval, 2-alternative forced-choice (2I2AFC) procedure is used throughout. The two signals presented in the two observation intervals have similar fundamental frequencies and, hence, similar residue pitches, and either have components of the same harmonic number, which we may call "coincident" components, or they have no common harmonics; that is, they have "noncoincident" components. In the former case, component frequency comparisons are possible, and performance need not be mediated by a residue pitch discrimination. In the latter case, component frequency comparisons are impossible.

IV. GENERAL EXPERIMENTAL METHODS

A. Stimuli

The stimuli were: (1) two-component harmonic complexes; (2) the individual components of a two-component complex; and (3) a reference stimulus with a strong and unambiguous pitch. The fundamental frequency of all the complexes was approximately 200 Hz. The two-component complexes contained either the 4th and 5th, or the 5th and 6th harmonics, while the reference signal contained the 1st, 2nd, 7th, 8th, and 9th harmonics. Experiments 1 and 3 employed two-component complexes and their individual components. In these experiments, the fundamental frequencies of the two stimuli given on each trial were equally spaced about 200 Hz. Experiments 2 and 4 employed the five-component reference signal and a two-component test signal. The reference signal had a constant fundamental frequency, and only the fundamental frequency of the two-component test signal was varied.

The complex signals were generated with components of equal amplitude added in sine phase. A low-pass-filtered masking noise was presented with the two-component complexes to mask any auditory distortion products below the frequencies of the components. The noise filter cutoff frequency was 613 Hz, the filter slope was 96 dB/oct, and the noise power density below the cutoff frequency was 20 dB below the level of the components. The stimuli were all 100 ms in half-power duration, and were gated with an exponential rise and fall time of 12.5 ms. Figure 1 shows the time course of the discrimination tasks.

B. Apparatus

The stimuli were computer generated (PDP8/e), played through 10 bit DACs, and recorded onto tape (Revox A77) after filtering (Kemo VBF/3) and gating (Grason–Stadler 829E). The tapes were replayed to the observers at twice the recording speed. The effective sampling rate of the signals at replay was 16 kHz, with low-pass filtering effectively at 3 kHz. A wave analysis of the recorded signals showed no distortion products less than 48 dB below the desired com-

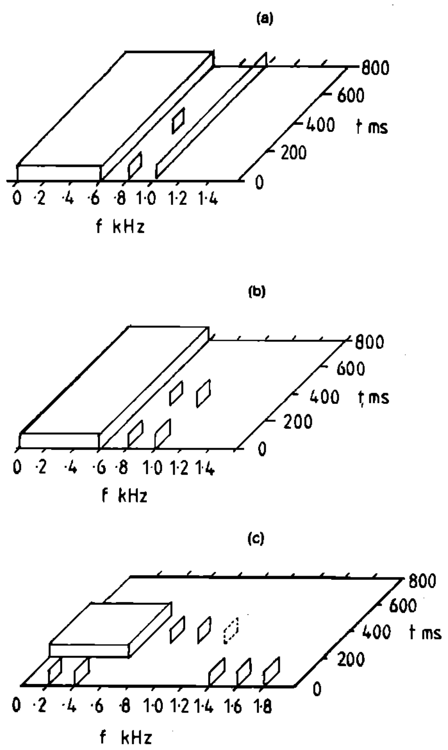

FIG. 1. The time course of the discrimination tasks; the gating envelope is not shown. The rectangular blocks represent the low-pass-filtered noise. Panel (a) illustrates the component frequency discrimination task; the test component is here shown at 800 Hz and the continuous second component is shown at 1000 Hz. Panel (b) shows the coincident component discrimination task; the components are at 800 and 1000 Hz, and both differ in frequency between the two intervals. Panel (c) shows noncoincident component discrimination; the first signal is the reference signal, with components at 200, 400, 1400, 1600, and 1800 Hz, and the second interval contains components at the 4th and 5th, or the 5th and 6th harmonics of 200 ± Δf Hz, embedded in a burst of noise.

ponents. The speed stability of the tape recorder over a record–replay cycle was better than 0.07% (peak reading). The observer was seated in a sound-insulated booth (Amplivox) and heard the stimuli through headphones (Rogers Ravensbrook) driven by a Mullard Unilex amplifier. The frequency response of the system as measured with an artificial ear (B & K type 4153 coupler, with type 4134 microphone, and type 2607 measuring amplifier) was flat within ± 2 dB between 200 and 2500 Hz. Noise was derived from a Dawe type 419C white noise generator, and filtered by a Kemo VBF/3 and a Dawe type 1462A filter in series.

C. Observers

Two observers were used. Each had normal hearing and was an active amateur musician. One observer was the author; the other was paid for his services.

D. Procedure

The method of constant stimuli was used; the two 100-ms observation intervals were separated by a 200-ms interval. The intertrial interval was about 3 s. Trials were presented in blocks with Δf constant; Δf was either increased or decreased monotonically throughout a session of eight blocks. The order of different experimental conditions was approximately counterbalanced, both within and between observers, over the days of testing.

The observer was instructed to judge the pitch of the tone in the first observation interval as "high" or "low" with respect to the pitch of the signal in the second interval, and wrote his response on an answer sheet. No trial-by-trial feedback was provided, but the observers were told of their general performance level when they next attended. The observers had at least 20 h of prior experience of pitch discrimination tasks, including a residue pitch discrimination task.

E. Analysis of results

The observers' performance was recorded in the form of two-tailed psychometric functions, showing the probability of the response "high" as a function of the relative frequency difference between the signals in the first and second intervals. These functions are summarized by DLs derived from the slopes, which were estimated by a Probit analysis (Finney, 1973); details of this procedure are given in Appendix B.

V. EXPERIMENT 1: DISCRIMINATION OF COMPONENT AND FUNDAMENTAL FREQUENCY OF COMPLEXES WITH COINCIDENT COMPONENTS

The first experiment addresses the question of whether the relative discriminability of the fundamental frequency of a two-component harmonic signal is better than that of the frequencies of the components of the signal. The varied parameter was the level of the components.

A. Stimuli and procedure

The two stimuli were composed of the 4th and 5th harmonics of 200 Hz. The presentation level of each component in different sessions was 15, 20, 30, 40, or 50 dB SL.[1] Trials were presented in blocks of 50 with Δf constant. The relative fundamental frequency differences were 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.008, and 0.012, except for the 15- and 50-dB component frequency discrimination data, where they were 0.002, 0.004, 0.006, 0.008, 0.012, 0.016, and 0.024. The psychometric functions were based on at least 100 trials per point.

B. Results and discussion

The psychometric functions are shown in Fig. 2. The Probit analysis showed that the observed functions from both tasks and both observers are generally well fitted by normal ogives on a linear frequency scale. The 50% points of the functions all lie very close to a zero frequency difference, indicating a negligible bias.

The DLs estimated from these functions (see Appendix B) are summarized in Table I. Figure 3 shows these data as a function of SL. The effect of SL is nonmonotonic, but consistent within and between each observer. The relative DLs decrease, as expected, between 15 and 30 dB SL, but start to rise again above 30 or 40 dB SL. This latter effect is probably due to an increased upward spread of masking from the low-pass noise at these higher levels (Weber, 1977); a similar increase in the frequency DL may be seen in Henning's (1967)

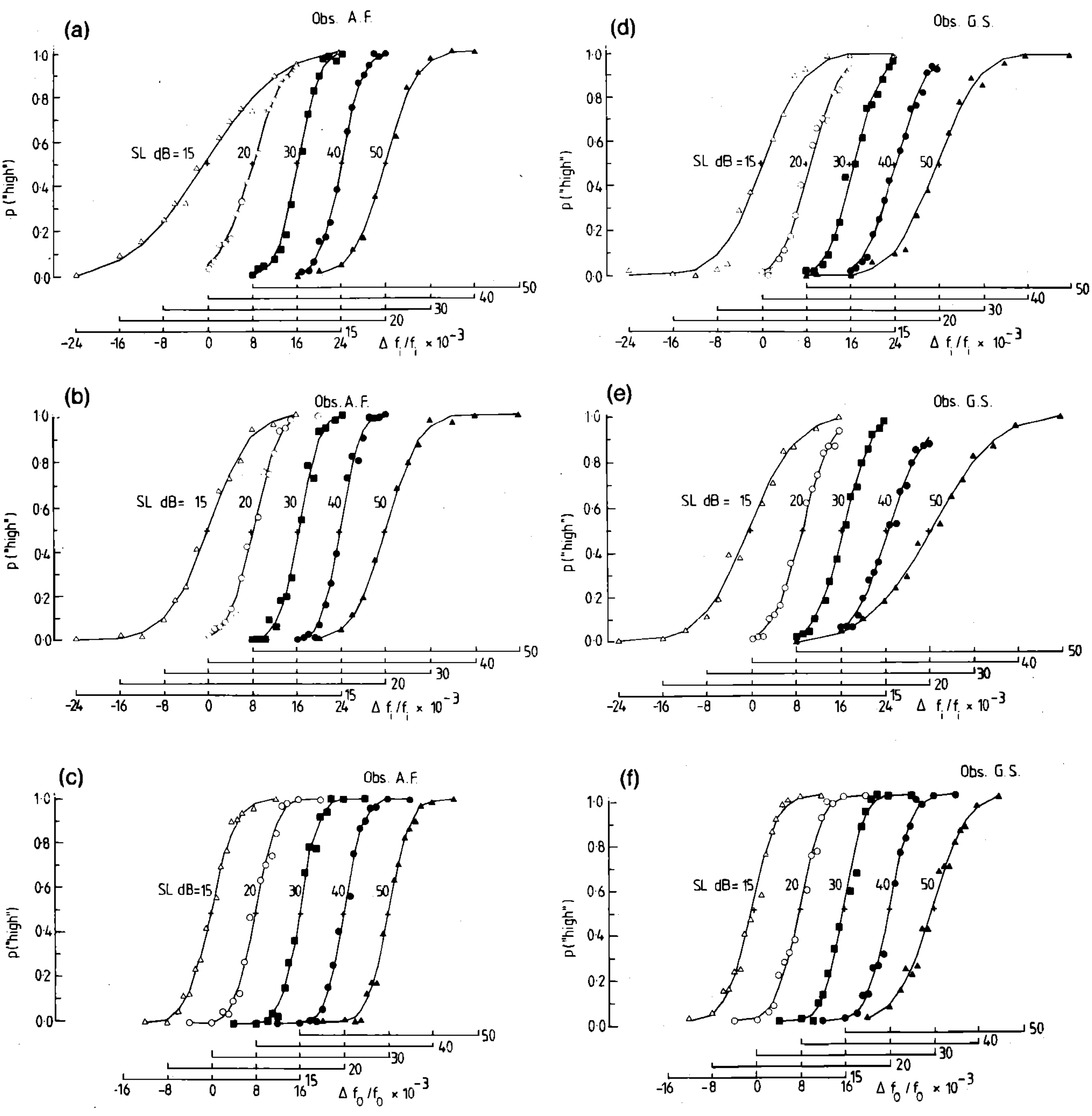

FIG. 2. Psychometric functions from experiment 1: component frequency discrimination, and complex signal discrimination with coincident components. The ordinate shows the probability of the response that the first signal is judged higher in pitch as a function of the relative frequency difference $\Delta f/f$ between the first and second signals, which is shown on the abscissae. The frequency-difference scale here and in the subsequent similar figures is displaced horizontally for each SL; each function and abscissa is labeled with the SL parameter. The curves are drawn through the best-fitting normal ogives given by a Probit analysis of the functions (see Appendix B). Panels (a) and (d) show the functions for the 800-Hz component, panels (b) and (e) show those for the 1000-Hz component, and panels (c) and (f) show the functions for the complex signal.

results at higher signal levels. The effect of SL appears to be greater for component frequency discrimination than for complex signal discrimination.

The finding of crucial interest is that the observed DLs for the complex signal are, in nine out of ten cases, smaller than the smaller relative DL of the component frequencies, and in eight out of the ten cases, this difference exceeds two standard errors. The observed DLs are close to the prediction of the multiple-band model and differ from the prediction by less than two standard errors in six out of the ten cases. The results thus support the multiple-band model. Since the relative DL for the complex could arise either from component frequency comparisons or from residue pitch comparisons, the data do not, in themselves, address the oth-

TABLE I. Relative DLs × 10^3 and standard errors for component frequency and coincident complex discrimination in experiment 1. The signal is indicated by harmonic numbers.

		15		20		30		40		50	
	SL dB	DL	SE	DL	SE	DL	SE	DL	SE	DL	SE
Obs AF	Harmonic number										
	4	7.56	0.37	3.50	0.11	2.26	0.09	2.27	0.09	3.36	0.17
	5	4.47	0.22	2.63	0.10	2.36	0.17	2.01	0.08	3.49	0.18
	4 + 5	2.72	0.12	2.31	0.10	2.00	0.09	2.03	0.09	2.36	0.10
Obs GS	Harmonic number										
	4	3.72	0.27	3.05	0.11	2.78	0.10	3.23	0.11	4.51	0.22
	5	4.98	0.24	3.09	0.13	2.92	0.12	3.83	0.14	6.63	0.40
	4 + 5	2.80	0.14	2.52	0.09	2.18	0.09	2.40	0.08	3.73	0.14

er issues of interest here. To pursue these issues, experiment 2 uses a task that prevents the use of component frequency comparisons.

VI. EXPERIMENT 2: DISCRIMINATION OF FUNDAMENTAL FREQUENCY WITH NONCOINCIDENT COMPONENTS

The second experiment required pitch discrimination between a pair of complex signals with noncoincident components. The major question is whether pitch discriminability with these signals is equivalent to that observed with coincident components. If performance is similar in both tasks, then both may be performed on the basis of residue pitch differences. If discriminability is poorer in the present task, it may be that discrimination of signals having coincident components is performed by reference to component frequency comparisons, and that residue pitch comparisons are used only when comparisons of individual components are impossible. As in experiment 1, the varied parameter was signal level.

A. Stimuli and procedure

A trial was composed of the reference signal described earlier followed by the test signal. The test signal contained either the 4th and 5th or the 5th and 6th harmonics of 200 ± Δf Hz. The reference signal was presented in a silent background with equal intensity components at a level equivalent to 40 dB SL at 1000 Hz. The test signal was presented with its components at 15, 20, 30, 40, or 50 dB SL together with the low-pass-filtered noise used in experiment 1. The noise was gated on 100 ms before the onset of the test signal, and off 100 ms after its offset, with a rise and fall time of 25 ms. The inclusion of two test signals within a block of trials was intended to encourage the use of fundamental frequency differences rather than between-trial comparisons of component frequencies; only the signal containing the 4th and 5th harmonics is comparable with that used in experiment 1.

Trials with a constant Δf were presented in blocks of 100, containing equal numbers of each of the two test signals. Eight blocks were given in a 1-h session with $\Delta f / f$ of 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.008, and 0.012. The observers received at least four sessions at each SL.

B. Results and discussion

The psychometric functions are shown in Fig. 4. The observed functions do not deviate systematically from normal ogives. The 50% points of the functions are close to a zero frequency difference for observer AF except for the test signal containing the 5th and 6th harmonics at 15 and 20 dB SL. For observer GS, the 50% point is clearly dependent on

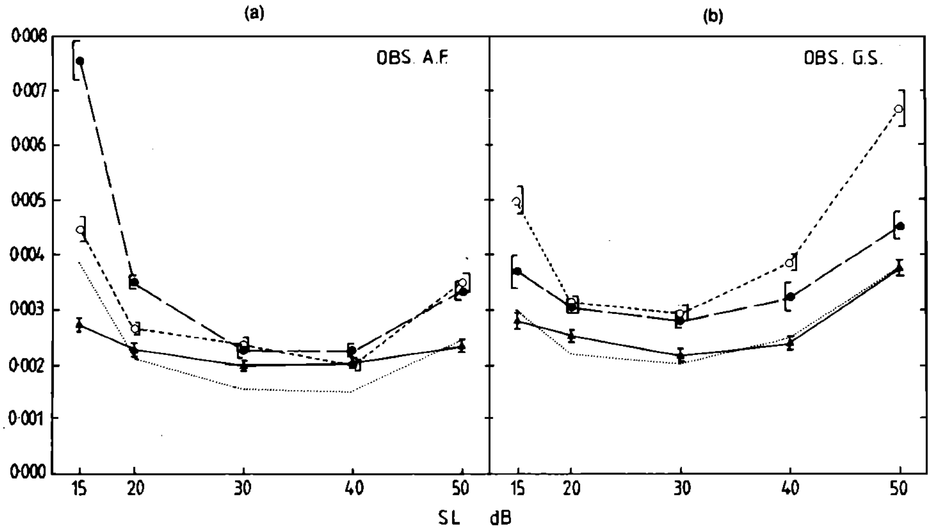

FIG. 3. Relative DLs as a function of SL for component frequency discrimination and complex signal discrimination with coincident components in experiment 1. Component frequency DLs are shown as filled circles (800 Hz) and open circles (1000 Hz), the complex signal DL as filled triangles, and the predicted complex DL as the dotted line. The error bars show ± 1 standard error.

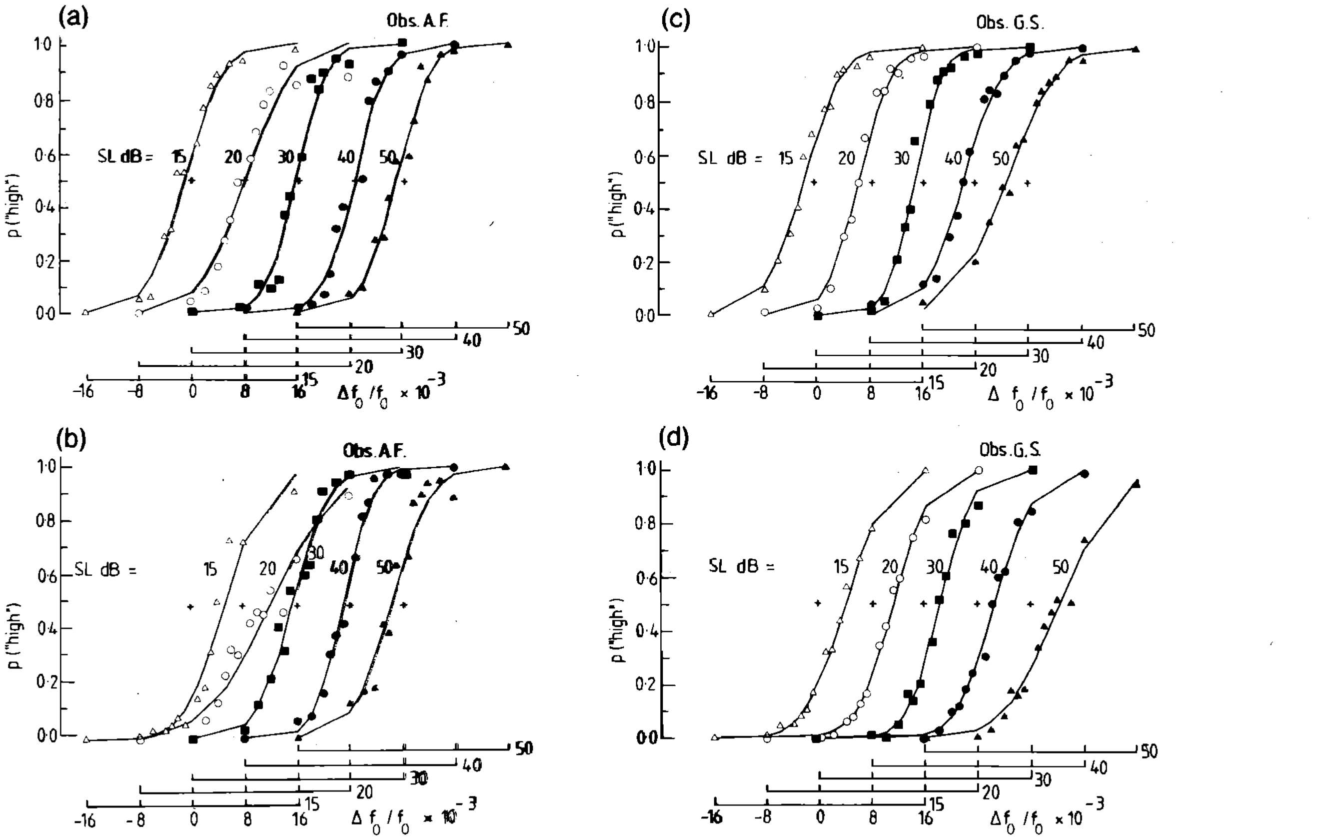

FIG. 4. Psychometric functions from experiment 2: noncoincident component complex discrimination with SL as the parameter. Axes as in Fig. 2. Panels (a) and (c) show the functions for the test signal comprising 800 and 1000 Hz, and panels (b) and (d) show the functions for 1000 and 1200 Hz.

the harmonic content of the test signal; for this observer, the reference signal was more likely to be judged "lower" than the test signal having the higher harmonics, and more likely to be judged "higher" than the test signal with lower harmonics. It is not clear whether this bias was due to a genuine pitch difference between the two test signals, or to a response bias related to the frequencies of the components. The main concern here, of course, is with the slopes of the functions from which the DLs are estimated.

Relative DLs were derived, as before, from the slopes of the psychometric functions (see Appendix B), and are shown in Table II. Figure 5 shows the relative DLs as a function of SL together with the complex signal DLs from experiment 1.

The DLs obtained at 20 dB from observer AF are elevated by unusually large between-session variability, and no significance is attached to this increase. If these points are neglected, the DLs show a comparable nonmonotonic dependence on SL to those found in experiment 1. In general, the pitch of the test signal containing the 4th and 5th harmonics was more discriminable than that of the signal containing the 5th and 6th harmonics, but this difference was slight.

The main result of the experiment is that the DLs observed here are consistently and significantly larger than those found for the same signal in the coincident discrimination task of experiment 1. This suggests that residue pitch differences are used only in the noncoincident component discrimination task, and not in the coincident component task. The difference between the coincident and the noncoincident complex DLs is consistent with an additional internal noise process, so that Goldstein's assumption that the precision of residue pitch is limited by an internal noise component may be correct. After data on the effects of signal-to-noise ratio on pitch discrimination have also been presented,

TABLE II. Relative DLs × 10^3 obtained from the noncoincident complex discrimination task of experiment 2.

		15		20		30		40		50	
	SL dB	DL	SE	DL	SE	DL	SE	DL	SE	DL	SE
Obs AF	Harmonic number										
	4 + 5	4.39	0.35	4.44	0.68	3.42	0.40	3.70	0.33	3.86	0.24
	5 + 6	4.79	0.50	5.57	0.75	4.18	0.26	3.17	0.17	4.56	0.54
Obs GS	Harmonic number										
	4 + 5	4.32	0.23	3.49	0.28	2.74	0.18	4.32	0.23	5.86	0.32
	5 + 6	4.59	0.26	4.02	0.23	3.57	0.26	4.29	0.23	6.19	0.42

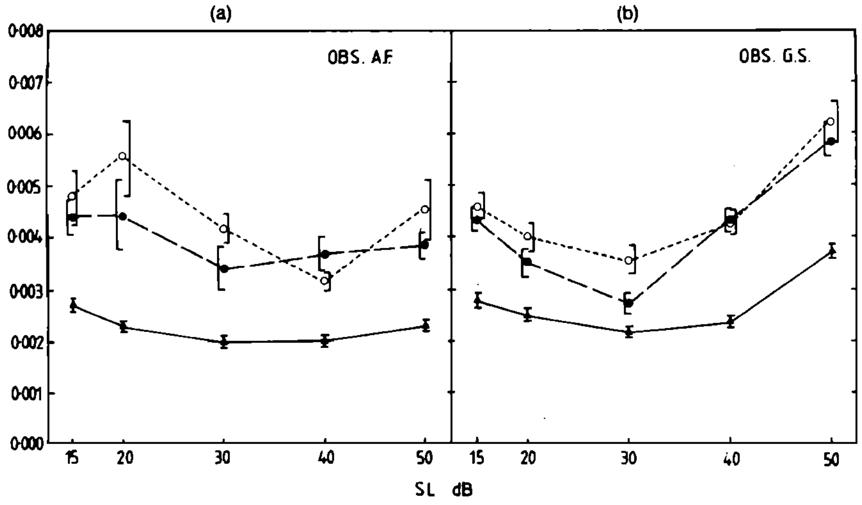

FIG. 5. Relative DLs as a function of SL from the coincident and noncoincident component complex discrimination tasks from experiments 1 and 2. Noncoincident DLs are shown as filled circles (800 and 1000 Hz) and empty circles (1000 and 1200 Hz); the coincident DLs are shown as solid triangles.

a quantitative test of this explanation will be made, which will also provide an estimate of the value of the internal noise contribution.

In order to achieve a greater range of discrimination performance, which allows a stronger test of the internal noise assumption, similar experiments were performed in which the signals were presented at a constant level in a variable signal-to-noise ratio.

VII. EXPERIMENTS 3 AND 4: DISCRIMINATION OF COMPONENT FREQUENCY AND FUNDAMENTAL FREQUENCY AS A FUNCTION OF SIGNAL-TO-NOISE RATIO

A. Stimuli and procedure

These two experiments replicate experiments 1 and 2, respectively, except that the test signals were always presented with each component at a constant SL of 30 dB, and a

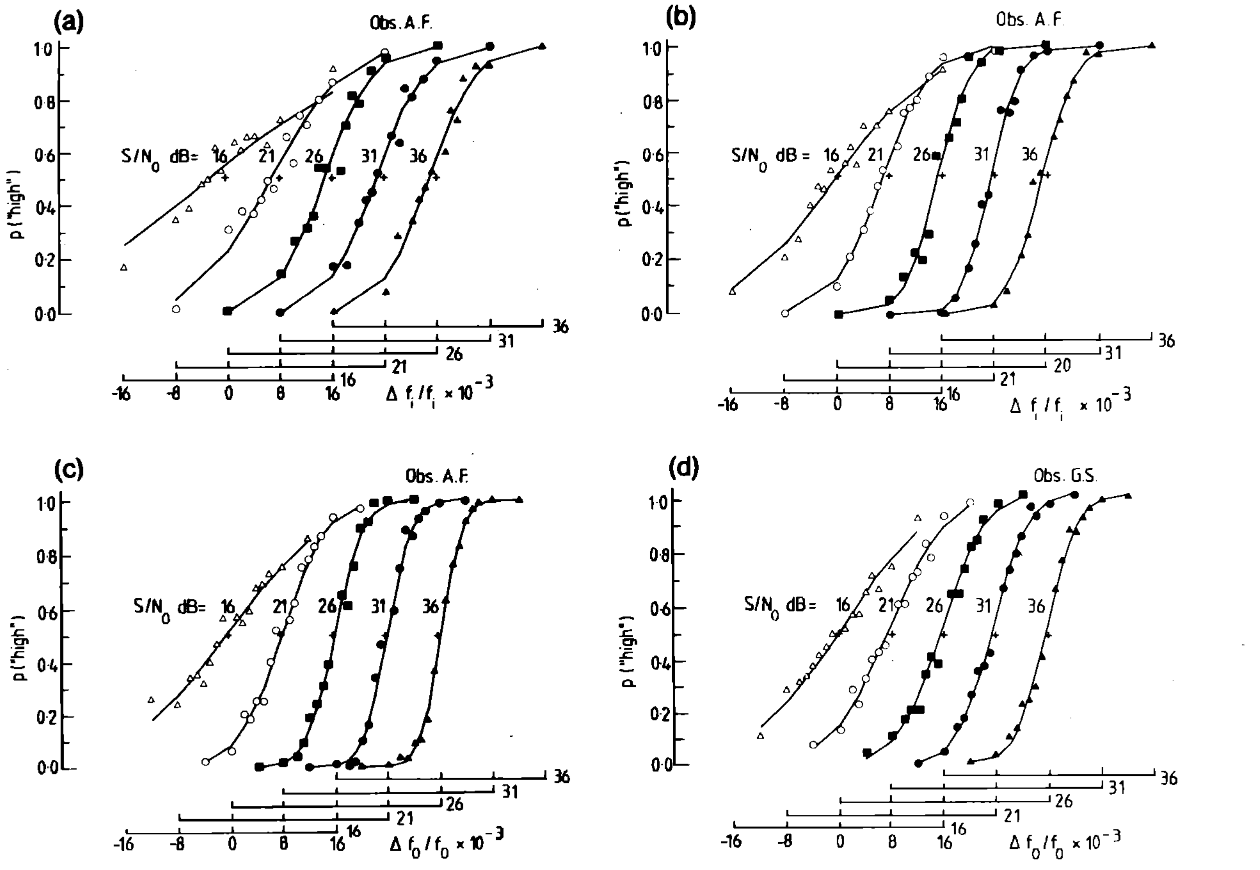

FIG. 6. Psychometric functions from experiment 3: component frequency discrimination, and complex signal discrimination with coincident components with S/N_0 as a parameter. Panels (a) and (b) show the functions for component frequency discrimination at 800 and 1000 Hz, respectively, and panels (c) and (d) show the functions for discrimination of fundamental frequency.

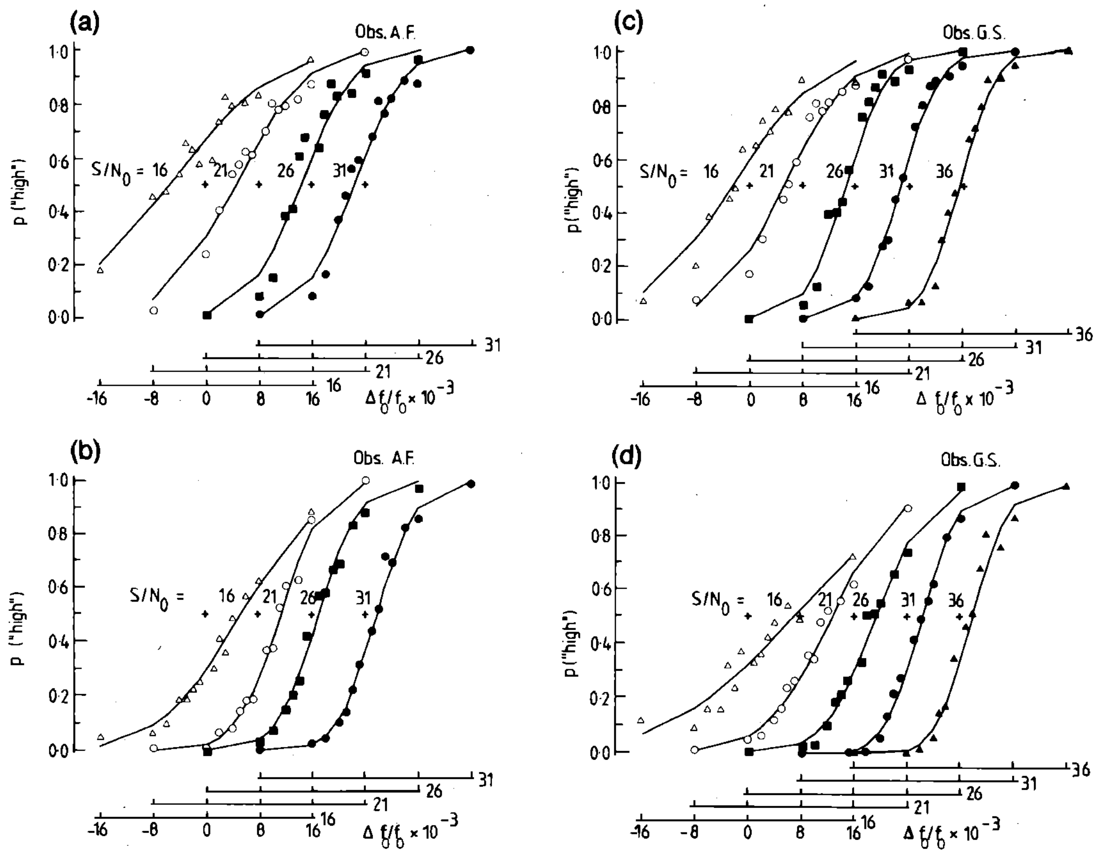

FIG. 7. Psychometric functions from experiment 4: noncoincident component complex discrimination. Panels (a) and (c) show the functions for the test signal comprising 800 and 1000 Hz, and panels (b) and (d) show the functions for the signal comprising 1000 and 1200 Hz.

broadband white noise of variable level was added to the low-pass masking noise. The white noise had a spectral density between 16, 21, 26, 31, or 36 dB below that of the signal components. The procedure was as before, except that only the author was able to participate in the component frequency discrimination task of experiment 3.

B. Results and discussion

The psychometric functions are shown in Figs. 6 and 7. All the functions are reasonably well fitted by the theoretical ogives. The functions from the component frequency and coincident tasks show minimal bias, but, in the noncoincident task, both observers show a consistent bias towards judging the test signal having the higher frequency components as higher in pitch, and vice versa.

The DLs were again derived from the slopes of the psychometric functions; these data are summarized in Table III. Figure 8 shows the data from observer AF in the component frequency discrimination task and the coincident component complex discrimination task together with the prediction for the latter task from the multiple-band model. The DLs all show a monotonically decreasing dependence on S/N$_0$. As the relative DL for the 4th harmonic is rather larger than that for the 5th, the prediction of the multiple band model is not very different from the smaller component frequency DL. Nevertheless, the observed DL for the coincident task is less than the smaller component frequency DL, where S/N$_0$ is 21 dB or more. The relative DLs from the noncoincident component complex in experiment 4 are shown in Fig. 9. The effect of S/N$_0$ is comparable with exper-

TABLE III. Relative DLs × 10^{-3} obtained from experiments 3 and 4. The subscripts "*c*" and "*n*" indicate the coincident and noncoincident tasks.

	S/N$_0$ dB	16 DL	SE	21 DL	SE	26 DL	SE	31 DL	SE	36 DL	SE
Obs AF	Harmonic number										
	4	12.04	0.88	6.49	0.39	4.40	0.24	4.36	0.24	4.14	0.22
	5	8.60	0.51	4.22	0.23	2.91	0.15	2.59	0.13	2.94	0.15
	$4+5_c$	8.88	0.64	4.04	0.18	2.61	0.11	2.28	0.10	1.97	0.09
	$4+5_n$	12.68	0.65	8.48	0.44	6.16	0.65	5.92	0.45	...	...
	$5+6_n$	9.89	0.50	5.19	0.26	4.82	0.33	4.59	0.23	...	...
Obs GS	Harmonic number										
	$4+5_c$	7.83	0.51	5.06	0.20	3.84	0.17	3.15	0.14	2.78	0.12
	$4+5_n$	10.38	0.80	7.87	0.55	4.64	0.47	4.37	0.30	3.94	0.25
	$5+6_n$	15.12	1.49	7.84	0.41	5.96	0.30	4.05	0.22	3.83	0.33

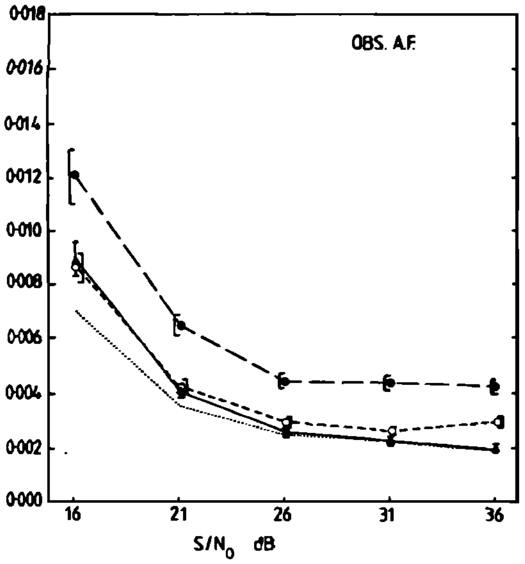

FIG. 8. Relative DLs as a function of S/N$_0$ from the component frequency and coincident component discrimination tasks from experiment 3. Data from observer AF only. The component DLs are shown as solid circles (800 Hz) and open circles (1000 Hz), and the fundamental frequency DL is shown as solid triangles; the multiple-band prediction for the fundamental frequency DL is given by the dotted line.

iment 3, but the noncoincident DLs are consistently larger than those from the coincident task. The present experiments replicate the finding from experiments 1 and 2; the noncoincident DLs exceed those for the coincident component complex, and the differences between the two sets of data are comparable in both cases. Hence, these data again support the idea that only the noncoincident component discrimination task involves the use of residue pitch and also support Goldstein's internal noise assumption. The observed DLs show a monotonic dependence on signal-to-noise ratio, but the noncoincident DLs fall only very slightly as the signal-to-noise ratio increases beyond 21 dB, although the coincident DLs continue to decrease. This behavior is characteristic of a noise-limited process.

VIII. GENERAL DISCUSSION

There may be alternative explanations of the differences between the performance observed in the coincident and noncoincident tasks. The use of pulsed rather than continuous masking noise in experiments 2 and 4 might be important (Weir *et al.*, 1977b), but the signals used here are 100 ms long and well above threshold, and no major effect is expected.

Perhaps more importantly, in the noncoincident task, there were two possible test signals on a trial, composed of two of three possible harmonics, while the same two-component test signal always occurred in the coincident task. This difference would lead to a degree of frequency uncertainty in the noncoincident task (see Appendix A), which might increase the DL in the noncoincident condition. This possibility is examined below.

If an added internal noise is responsible for the observed effect, then the equivalent internal variance may be estimated as

$$\sigma^2(\text{int}) = \sigma_n^2 - u(\sigma_c^2), \tag{4}$$

where σ_n is the relative DL for the noncoincident case and σ_c is that for the coincident case. As there is an uncertainty difference between the two tasks which may also act to increase the DL, a multiplicative term u is also included, representing the ratio of uncertainty between the noncoincident and the coincident discrimination tasks. The additive and multiplicative terms may be estimated from the fitting of a structural relation (Kendall and Stuart, 1961) to the relation between σ_n^2 and σ_c^2. The results of this analysis are given in Table IV. The best fitting values of u were 1.92 and 1.78 for observers AF and GS, respectively; both values are significantly greater than 1, indicating that frequency uncertainty probably makes a contribution to the performance difference. The best-fitting internal noise parameters were 0.0038 and 0.0019, respectively, corresponding to internal limits for the complex pitch DL of 0.76 and 0.34 Hz, respectively, at 200 Hz. The estimated $\sigma(\text{int})$ arises from both components of the test signal, and from Eq. (2), Goldstein's parameter $\sigma^2(f_i)/f_i$ contains an average internal contribution of

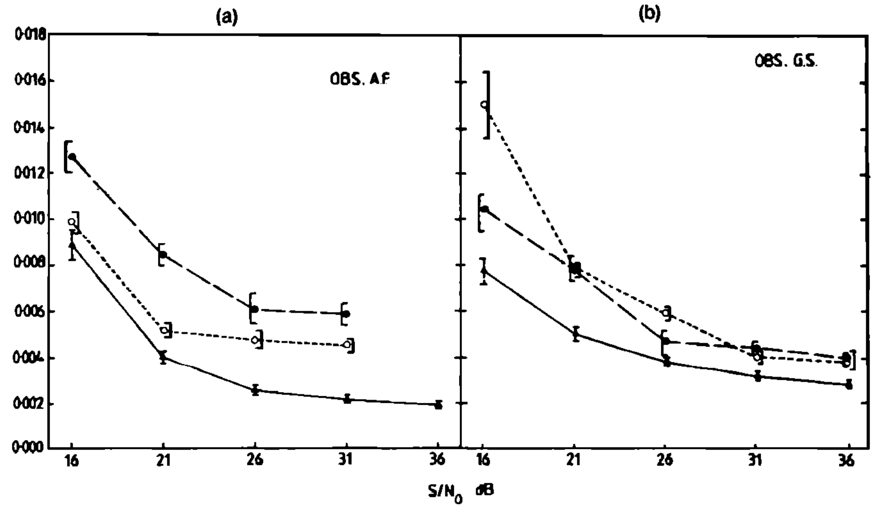

FIG. 9. Relative DLs as a function of S/N$_0$ from the coincident and noncoincident component complex signal discrimination tasks from experiments 3 and 4. The DLs for the 800- and 1000-Hz signal are shown as filled circles, and the DLs for the 1000- and 1200-Hz signal as open circles; the relative DLs from the coincident task are shown by the filled triangles.

TABLE IV. Structural relation analysis of dependence of σ_n^2 on σ_c^2. Data from experiments 1–4. u is the slope parameter, and σ(int) the square root of the intercept of the best fitting straight line. r is the correlation between σ_n^2 and σ_c^2. The t statistic was computed against a theoretical slope of 1 corresponding to no uncertainty effect. λ is the variance ratio of σ_n^2 and σ_c^2, and was estimated from the data standard errors.

Observer	u	r	σ(int)	t	df	p	λ
AF	1.924	0.964	0.00377	6.87	2	<0.05	40.0
GS	1.776	0.979	0.00193	8.37	2	<0.01	10.0

$2\sigma^2$(int), giving $\sigma(\text{int}_i)$ of 0.0054 and 0.0027 for the two observers, respectively. These values correspond to an effective limit to the precision of estimation of frequency of 5.4 and 2.7 Hz, respectively, for a component at 1000 Hz. These results are consistent with Goldstein's internal noise assumption, although the estimated noise parameters are somewhat smaller than those estimated from musical interval identification tasks (Goldstein, 1973; Goldstein *et al.*, 1978; Houtsma, 1979), which would suggest a component frequency precision limit of 10 Hz or more at 1000 Hz.

Finally, it is of interest to compare the effects of S/N_0 observed here with Houtgast's (1976) report that a fundamental frequency difference was more accurately detected in a low S/N ratio than in quiet. The present data show that, as in pure-tone discrimination (e.g., Henning, 1967), pitch discrimination improves as S/N increases. The DLs measured here at S/N_0 above 26 dB are between 0.002 and 0.006; this is considerably less than the relative fundamental frequency difference of 0.030 employed by Houtgast, and it appears that Houtgast's intriguing result does not reflect a limit of auditory discrimination ability.

IX. CONCLUSIONS

The major finding of the studies reported here is that pitch discrimination with complex signals depends not only on the factors affecting the discriminability of simple signals, but also upon whether observers are able to make comparisons between the pitches of components which are common to a pair of signals. Where comparisons between component frequencies are impossible, an observer must compare some second-order percept, such as the residue pitch of the signals. This result implies that experiments, in which signals with common components are compared, may not teach us anything about the perception of residue pitch. In particular, the DLs reported by Walliser (1969) and Moore *et al.* (1984) are probably not DLs for residue pitch, at least where one or more components can be resolved, and they are probably the result of statistical summation over cues from several resolved components. This study also confirms Goldstein's (1973) hypothesis that the precision of residue pitch perception is limited by internal variability, and refutes the contrary claim of Moore *et al.* (1984), which was based upon discrimination performance with coincident components.

The present experiments provide a direct estimate of the proposed variability, whereas previous approaches (Goldstein, 1973; Goldstein *et al.*, 1978; Houtsma, 1979) provided only indirect estimates from parameter fitting. The present estimates of the internal variability, based on data from signals with components between 800 and 1600 Hz, correspond to a relative standard deviation between 0.0027 and 0.0054. These estimates are smaller than those made by Goldstein and his colleagues, but it should be noted that factors such as frequency uncertainty may act to increase the effective internal variability; the identification experiments from which Goldstein's estimates are derived, by their nature, show greater frequency uncertainty than the discrimination experiments presented here. Although Goldstein is the only theorist who has explicitly discussed internal variability, the generally similar models described by Terhardt (1970, 1974) and Wightman (1973) could be readily modified to include this component.

A third finding of the present studies is that where observers can make pitch comparisons between common components, a multiple-band statistical summation model developed from a similar model for detection (Green, 1958) provides a good account of the relation between the discriminability of frequencies of the individual components and the discriminability of the pitch of the complex signal. This confirms the finding reported by Moore *et al.* (1984) who employed a somewhat different experimental method.

X. SUMMARY

(1) The results confirm the applicability of a multiple-band theory to frequency discrimination of complex tones. The theory applies directly where coincident component frequency comparisons may be made. This confirms the finding of Moore *et al.* (1984), but, contrary to their claim, such discrimination does not depend on the precision of a residue pitch.

(2) Where coincident frequency comparisons cannot be made, but residue pitch comparisons may be made, there is an internal contribution to the variability of pitch precision. This conclusion is in contradiction with that reached by Moore *et al.* (1984), who failed to find evidence for an internal noise, but only considered a task allowing coincident frequency comparisons.

ACKNOWLEDGMENTS

The author thanks Dr. M. Treisman for his continued support and encouragement, and is indebted to Dr. G. B. Henning and two anonymous reviewers for many helpful comments on a previous draft of this manuscript. This research was partly supported by a studentship funded by the Science Research Council of Great Britain. Support from the Medical Research Council is also acknowledged. The experiments were earlier described in a D. Phil. thesis submitted to the University of Oxford. Parts of this paper were presented at the July 1984 joint meeting of the Experimental Psychology Society and the Netherlands Psychonomic Foundation.

APPENDIX A: A MULTIPLE-BAND MODEL FOR PITCH DISCRIMINATION

1. A simple model

Consider a conventional 2I2AFC pitch discrimination task for a complex signal. If N frequency differences of the same sign occur simultaneously in N independent critical bands, an optimal multiple-band probability summation process may combine the evidence over channels. As Green (1958) has shown in an analogous detection task, d'_c, the index of the detectability of at least one of the N frequency differences is given by

$$d'_c = \left(\sum_{i=1}^{N} d_i'^2 \right)^{1/2}, \tag{A1}$$

where d'_i is the detectability index for the ith component frequency difference.

Now d'_i is inversely proportional to the standard deviation of the evidence for a frequency difference, which is the frequency DL $\delta(f_i)$, and, if the frequencies of all the components of a complex signal change in the same direction, the relative DL for the pitch change of the complex signal $\delta(f_c)/f_c$ must be given by

$$\left(\frac{f_c}{\delta(f_c)} \right)^2 = \sum_{i=1}^{N} \left(\frac{f_i}{\delta(f_i)} \right)^2, \tag{A2}$$

which, equating $\delta(f_c)$ with $\delta(f_0)$, gives Eq. (1).

2. The effect of frequency uncertainty

There is considerable evidence from auditory detection studies that multiple-band processes are subject to an effect of frequency uncertainty (Tanner *et al.*, 1956; Veniar, 1958a,b; Creelman, 1960; Greenberg and Larkin, 1968) and that frequency uncertainty also affects frequency discrimination of complex signals comprising temporally spaced components (Spiegel and Watson, 1981) and discrimination of complex signals with simultaneous components (Faulkner, 1982). However, Green's (1958) model does not predict such an effect, since channels which receive only noise give $d'_i = 0$ and, hence do not affect d'_c. It is possible to construct an alternative multiple-band model in which the decision statistic is given by

$$d'_c = \sum_{i=1}^{N} \mu(x_i) \Bigg/ \left(\sum_{i=1}^{N} \sigma^2(x_i) \right)^{1/2}, \tag{A3}$$

where x_i is the normally distributed evidence variable in the ith channel, N channels respond to signal + noise, and M–N channels ($M<N$) respond to noise only. This model does produce an effect of uncertainty, since the inclusion of evidence from channels with no signal will decrease d'_c. However, the model is non-optimal, and, if different channels have different d'_is, the overall d'_c will be biased towards the smaller d'_is, whereas Green's model will always give d'_c which is at least as large as the largest d'_i.

Frequency uncertainty effects may be better accounted for by the addition to Green's model of the assumption that observers have limited processing capability which may be distributed either to maximize sensitivity in a few channels, or to produce more moderate sensitivity in a larger number of channels.

APPENDIX B: ANALYSIS OF PSYCHOMETRIC FUNCTIONS

Each observed psychometric function was subjected to a Probit analysis (Finney, 1971), which estimates the normal ogive that best fits the observed function according to a maximum-likelihood criterion. It was assumed that the psychometric functions were normal ogives on a linear frequency difference scale, and this assumption was generally justified by the good fit of the theoretical curve. The best-fitting normal ogive is an estimate of the cumulated distribution of differences assumed to underlie discrimination performance. The reciprocal of the standardized slope of the ogive is then an estimate of the standard deviation of the distribution of differences σ_d which is composed of

$$\sigma_d^2 = \sigma_{S1}^2 + \sigma_{S2}^2, \tag{B1}$$

where σ_{Si}^2 is the variance of an internal representation of the frequency or pitch of the signal in the ith interval.

In experiments 1 and 3, similar signals were presented in the two observation intervals, and these may be assumed to contribute equally to σ_d^2, so that the DL, which is equivalent to σ_{Si}, may be estimated as $\sigma_d/\sqrt{2}$. In experiments 2 and 4, the reference and test signals presented in the two observation intervals differed in their composition and presentation conditions, and it is unreasonable to assume that $\sigma_{S1}^2 = \sigma_{S2}^2$. Hence, the relative DL for the test signal (S2) is derived by subtraction of σ_{S1}^2 from σ_d^2, where σ_{S1}^2 was taken as the DL estimated by the first method above from a 2I2AFC task with the reference signal (S1) presented in both intervals. The estimates of the relative DL for the reference signal are given in Table B 1.

Probit analysis allows the computation of standard errors and confidence limits for the estimated DLs based not only on the theoretical binomial variance, but also on any additional variability. Error bars based on an interval of ± 1 standard error are included in the graphs illustrating the estimated DLs. The standard error for the relative DL of the reference signal is included in Table B1.

TABLE B1. Relative DLs and standard errors of estimation for the pitch of the reference signal used in experiments 2 and 4.

Observer	DL	S.E.
AF	0.00181	0.00008
GS	0.00207	0.00006

---

[1] Zero dB SL was assumed to be roughly constant between 800 and 1400 Hz (Fletcher and Munson, 1933; Robinson and Dadson, 1956), and was determined as the intensity giving $d' = 1$ in a 2 interval forced-choice detection task with a 1-kHz, 100-ms pure tone. Zero dB SL was equivalent to 10 dB SPL for observer AF and 8 dB SPL for observer GS.

Boer, de, E. (1977). "Pitch theories unified," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 323–334.

Creelman, C.D. (1960). "Detection of signals of uncertain frequency," J. Acoust. Soc. Am. 32, 805–810.

Cudahy, E. A. (1975). "Fundamental pitch discrimination for monotic and dichotic two-tone complexes," J. Acoust. Soc. Am. Suppl. 1 57, S27.

Fastl, H., and Weinberger, M. (1981). "Frequency discrimination for pure and complex tones," Acustica **49**, 77–78.

Faulkner, A. (1982). "Models of the perception of the pitch of tonal complexes," D. Phil. thesis, University of Oxford.

Finney, D. J. (1973). *Probit Analysis* (Cambridge U. P., Cambridge, England), 3rd ed.

Flanagan, J. L., and Saslow, M.G. (1958). "Pitch discrimination for synthetic vowels," J. Acoust Soc. Am. **30**, 435–442.

Fletcher, H., and Munson, W. A. (1933). "Loudness, its definition, measurement, and calculation," J. Acoust. Soc. Am. **5**, 82–108.

Goldstein, J. L. (1973). "An optimum processor theory for the central formation of the pitch of complex tones," J. Acoust. Soc. Am. **54**, 1496–1516.

Goldstein, J. L., Gerson, A., Srulovicz, P., and Furst, M. (1978). "Verification of the optimal probabilistic basis of aural processing in pitch of complex tones," J. Acoust. Soc. Am. **63**, 486–497.

Green, D. M. (1958). "Detection of multiple component signals in noise," J. Acoust. Soc. Am. 30, 904–911.

Greenberg, G. Z., and Larkin, W. D. (1968). "Frequency-response characteristics of auditory observers detecting signals of a single frequency in noise: The probe signal method," J. Acoust. Soc. Am. **44**, 1513–1523.

Harris, J. D. (1952). "Pitch discrimination," J. Acoust. Soc. Am. **24**, 750–755.

Harris, J. D. (1966). "Masked DL for pitch memory," J. Acoust. Soc. Am. **40**, 43–46.

Helmholtz, von, H. L. F. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage fur die Theorie der Musik* (F. Veiweg & Sohn, Braunschweig, 1863), translated into English by A. J. Ellis, *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Dover, New York, 1954).

Henning, G. B. (1966). "Frequency discrimination for random amplitude tones," J. Acoust. Soc. Am. **39**, 336–339.

Henning, G. B. (1967). "Frequency discrimination in noise," J. Acoust. Soc. Am. **41**, 774–777.

Henning, G. B. (1970). "A comparison of the effects of signal duration on frequency and amplitude discrimination," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden, The Netherlands), pp. 350–361.

Henning, G. B., and Grosberg, S. L. (1968). "Effect of harmonic components on frequency discrimination," J. Acoust. Soc. Am. **44**, 1386–1389.

Horst, J. W., Ritsma, R. J., and Wit, H. P. (1984). "Frequency discrimination in quiet and in noise for signals with triangular spectral envelopes," J. Acoust. Soc. Am. **76**, 1067–1075.

Houtgust, T. (1976). "Subharmonic pitches of a pure tone at low S/N ratio," J. Acoust. Soc. Am. **60**, 405–409.

Houtsma, A. J. M. (1979). "Musical pitch of two-tone complexes and predictions by modern pitch theories," J. Acoust. Soc. Am. **66**, 87–89.

Kendall, M. G., and Stuart, A. (1961). *The Advanced Theory of Statistics, Vol. 2. Inference and Relationship* (Griffin, London).

Liang, C., and Chistovich, L. A. (1961). "Frequency-difference limens as a function of tonal duration," Sov. Phys. Abstr.: Acoust. **6**, 75–80.

Moore, B. C. J. (1973). "Frequency difference limens for short duration tones," J. Acoust. Soc. Am. **54**, 610–619.

Moore, B. C. J., Glasberg, B. R., and Shailer, M. J. (1984). "Frequency and intensity difference limens for harmonics within complex tones," J. Acoust. Soc. Am. **75**, 550–561.

Ritsma, R. J. (1963). "On pitch discrimination of residue tones," Int. Audiol. **2**, 34–37.

Robinson, D. W., and Dadson, R. S. (1956). "A redetermination of the equal-loudness relations for pure tones," Br. J. Appl. Phys. **7**, 166–181, 392–398.

Scheffers, M. T. M. (1984). "Discrimination of fundamental frequency of synthesized vowel sounds in a noise background," J. Acoust. Soc. Am. **76**, 428–434.

Schodder, G. F., and David, E. E., Jr. (1960). "Pitch discrimination of two-frequency complexes," J. Acoust. Soc. Am. **32**, 1426–1435.

Shower, E. G., and Biddulph, R. (1931). "Differential pitch sensitivity of the ear," J. Acoust. Soc. Am. **3**, 275–287.

Spiegel, M. F., and Watson, C. S. (1981). "Factors in the discrimination of tonal patterns. III. Frequency discrimination with components of well-learned patterns," J. Acoust. Soc. Am. **69**, 223–230.

Spiegel, M. F., and Watson, C. S. (1984). "Performance on frequency-discrimination tasks by musicians and nonmusicians," J. Acoust. Soc. Am. **76**, 1690–1695.

Srulovicz, P., and Goldstein, J. L. (1983). "A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum," J. Acoust. Soc. Am. **73**, 1266–1276.

Tanner, W. P., Jr., Swets, J. A., and Green, D. M. (1956). "Some general properties of the hearing mechanism," University of Michigan Electronic Defense Group, Technical Report No. 30.

Terhardt, E. (1970). "Frequency analysis and periodicity detection in the sensations of roughness and periodicity pitch," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden, The Netherlands), pp. 278–287.

Terhardt, E. (1974). "Pitch, consonance, and harmony," J. Acoust. Soc. Am. **55**, 1061–1069.

Turnbull, W. W. (1944). "Pitch discrimination as a function of tonal duration," J. Exp. Psychol. **34**, 302–316.

Veniar, F.A. (1958a). "Signal detection as a function of frequency ensemble. I," J. Acoust. Soc. Am. **30**, 1020–1024.

Veniar, F.A. (1958b). "Signal detection as a function of frequency ensemble. II," J. Acoust. Soc. Am. **30**, 1075–1078.

Walliser, K. (1969). "Zur Unterschiedsschwelle der Periodentonhohe," Acustica **21**, 329–336.

Weber, D. L. (1977). "Growth of masking and the auditory filter," J. Acoust. Soc. Am. **62**, 424–429.

Weir, C. C., Jesteadt, W., and Green, D. M. (1977a). "Frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **61**, 178–184.

Weir, C. C., Green, D. M., Hafter, E. R., and Burkhardt, S. (1977b). "Detection of a tone burst in continuous- and gated-noise maskers; Defects of signal frequency, duration, and masker level," J. Acoust. Soc. Am. **61**, 1298–1300.

Wightman, F. L. (1973). "The pattern-transformation model of pitch," J. Acoust. Soc. Am. **54**, 407–418.