

MODELS OF THE PERCEPTION OF THE PITCH OF TONAL COMPLEXES.

Andrew Faulkner.

New College.

Thesis submitted for the degree of D. Phil.,

University of Oxford,

Trinity Term, 1982.



MODELS OF THE PERCEPTION OF THE PITCH OF TONAL COMPLEXES

Andrew Faulkner, New College,

D. Phil. Thesis, University of Oxford.

Trinity Term, 1982.

ABSTRACT.

Theories of pitch perception, and the related literature, are reviewed, with special reference to the residue pitch of tonal complexes. A distinction is drawn between spectral-pattern theories, which propose that pitch is derived from independent internal estimates of component frequency, and periodicity theories, which propose that the pitch of tonal complexes may be derived from a residual periodicity resulting from an incomplete auditory frequency analysis. The Spectral-pattern models described by Goldstein (1973) and Wightman (1973b) are discussed in detail; computer simulation procedures, allowing the prediction of a probability density function for the estimated fundamental frequency of a tonal complex, are described for each. Contrasting predictions concerning the relation between component frequency discrimination and fundamental frequency discrimination for harmonic complexes are developed for the spectral-pattern and periodicity theories.

Component frequency and fundamental frequency discrimination by human observers was measured, under a variety of conditions, for a complex containing the 4th and 5th harmonics of 200 Hz. For a comparison of the pitches of two such complexes of slightly different frequency, the relative fundamental frequency difference limen was no larger, and generally smaller, than the smaller relative component frequency difference limen. Discrimination performance was interpreted in terms of a modified version of Goldstein's (1973) spectral-pattern theory, in which an internal noise limits discrimination between signals with non-coincident components. Further experiments investigated the identification of fundamental frequency by human observers as a function of the variability of internal estimates of component frequency as estimated from the discrimination experiments. The results of these experiments were consistent with the predictions of the spectral-pattern theories.

Component and fundamental frequency discrimination was also considered with respect to temporal and spatial theories of frequency discrimination. The results of the discrimination studies were consistent with a modified Energy Detection model, where discrimination is ultimately limited by an internal noise.

ACKNOWLEDGEMENTS.

The research reported in this thesis was carried out while the author held a three-year Research Studentship from the Science Research Council.

Acknowledgements are also due to Michel Treisman, who supervised this research, and read previous drafts of the thesis with great care and patience. I am grateful also for the assistance and support provided by Bruce Henning, and for the patient dedication of Geoffrey Sample, who spent many hundreds of hours acting as an observer. Thanks are due to Ken Ford, Peter Coughlin, and their staff, for administrative and technical assistance, to John Collins and Ivor Lloyd for their freely given time and advice on computing matters. My wife Dorothy helped in the preparation of figures and in proof-reading, and also gave continual support and encouragement throughout this project. Finally, I would like to thank the teaching staff of the University of Stirling Psychology Department, who taught me the excitement of psychological research, and much else besides.

INDEX TO FIGURES.

Figure 1.1	Following page 1-4
Figure 1.2	Following page 1-11
Figure 1.3	Following page 1-15
Figure 1.4	Following page 1-16
Figure 1.5	Following page 1-18
Figure 1.6	Following page 1-19
Figure 1.7	Following page 1-21
Figure 1.8	Following page 1-25
Figure 1.9	Following page 1-27
Figure 1.10	Following page 1-29
Figure 1.11	Following page 1-29
Figure 1.12	Following page 1-31
Figure 1.13	Following page 1-32
Figure 1.14	Following page 1-35
Figure 1.15	Following page 1-36
Figure 1.16	Following page 1-40
Figure 1.17	Following page 1-41
Figure 1.18	Following page 1-44
Figure 1.19	Following page 1-46
Figure 1.20	Following page 1-48
Figure 1.21	Following page 1-52

Figure 2.1	Following page 2-5
Figure 2.2	Following page 2-6
Figure 2.3	Following page 2-7
Figure 2.4	Following page 2-10
Figure 2.5	Following page 2-17
Figure 2.6	Following page 2-20
Figure 2.7	Following page 2-21
Figure 2.8	Following page 2-21
Figure 2.9	Following page 2-23
Figure 2.10	Following page 2-23
Figure 2.11	Following page 2-24
Figure 2.12	Following page 2-24
Figure 2.13	Following page 2-25
Figure 2.14	Following page 2-25

Figure 3.1	Following page 3-5
Figure 3.2	Following page 3-6
Figure 3.3	Following page 3-21
Figure 3.4	Following page 3-25
Figure 3.5	Following page 3-29
Figure 4.1	Following page 4-7
Figure 4.2	Following page 4-13
Figure 4.3	Following page 4-18
Figure 4.4	Following page 4-18
Figure 4.5	Following page 4-20
Figure 4.6	Following page 4-20

Figure 5.1	Following page 5-7
Figure 5.2	Following page 5-9
Figure 5.3	Following page 5-11
Figure 5.4	Following page 5-13

Figure 6.1	Following page 6-11
Figure 6.2	Following page 6-13
Figure 6.3	Following page 6-13
Figure 6.4	Following page 6-15
Figure 6.5	Following page 6-15
Figure 6.6	Following page 6-15
Figure 6.7	Following page 6-20
Figure 6.8	Following page 6-20
Figure 7.1	Following page 7-10
Figure 7.2	Following page 7-10
Figure 7.3	Following page 7-10
Figure 7.4	Following page 7-15
Figure 7.5	Following page 7-15
Figure 7.6	Following page 7-15
Figure 7.7	Following page 7-15
Figure 7.8	Following page 7-15
Figure 7.9	Following page 7-19
Figure 7.10	Following page 7-19
Figure 7.11	Following page 7-21
Figure 7.12	Following page 7-23
Figure 7.13	Following page 7-24
Figure 8.1	Following page 8-6
Figure 8.2	Following page 8-6
Figure 8.3	Following page 8-8
Figure 8.4	Following page 8-8
Figure 8.5	Following page 8-11
Figure 8.6	Following page 8-11
Figure 8.7	Following page 8-11
Figure 9.1	Following page 9-9
Figure 9.2	Following page 9-11
Figure 9.3	Following page 9-11
Figure 9.4	Following page 9-11
Figure 9.5	Following page 9-11
Figure 10.1	Following page 10-9
Figure 10.2	Following page 10-23
Figure 10.3	Following page 10-24
Figure 10.4	Following page 10-25
Figure 10.5	Following page 10-26
Figure 10.6	Following page 10-26
Figure 10.7	Following page 10-26
Figure 10.8	Following page 10-26
Figure 10.9	Following page 10-26
Figure 10.10	Following page 10-26

INDEX TO TABLES.

Table 1.1	Page 1-53
Table 4.1	Page 4-8
Table 4.2	Page 4-9
Table 4.3	Page 4-9
Table 4.4	Page 4-9
Table 4.5	Page 4-19
Table 4.6	Page 4-19
Table 5.1	Page 5-8
Table 5.2	Page 5-12
Table 6.1	Page 6-13
Table 6.2	Page 6-17
Table 6.3	Page 6-17
Table 6.4	Page 6-18
Table 7.1	Page 7-16
Table 7.2	Page 7-16
Table 7.3	Page 7-17
Table 7.4	Page 7-17
Table 7.5	Page 7-17
Table 7.6	Page 7-25
Table 8.1	Page 8-9
Table 8.2	Page 8-9
Table 9.1	Page 9-6
Table 9.2	Page 9-12
Table 9.3	Page 9-14
Table 9.4	Page 9-16
Table 10.1	Page 10.7
Table 10.2	Page 10.8
Table 10.3	Page 10.11
Table 10.4	Page 10.15

CONTENTS.

The thesis contains about 75,000 words.

Abstract.	-	-	-	-	-	-	i.
Acknowledgements.	-	-	-	-	-	-	ii.
Index to Figures.	-	-	-	-	-	-	iii.
Index to Tables.	-	-	-	-	-	-	v.
Extended Abstract.	-	-	-	-	-	-	1

Chapter 1 - A Review of the Perception of the Pitch of Tonal Complexes.

Abstract.	-	-	-	-	-	-	1-1
Introductory Remarks.	-	-	-	-	-	-	1-3
The Quantification of Pitch.	-	-	-	-	-	-	1-5
Psychological Theories of Pitch Perception.	-	-	-	-	-	-	1-6
An Early Theory, Ohm's Acoustical Law.	-	-	-	-	-	-	1-8
The Place Theory, a Physiological Basis for Ohm's Law.	-	-	-	-	-	-	1-9
Seebeck, Schouten, and the Pitch of the Residue.	-	-	-	-	-	-	1-10
The Response of the Place Theorists: Auditory Nonlinearity.	-	-	-	-	-	-	1-13
Frequency Analysis and Residual Periodicity.	-	-	-	-	-	-	1-16
The Perception of Residue Pitch as Spectral-Pattern	-	-	-	-	-	-	-
Recognition.	-	-	-	-	-	-	1-20
Residue Pitch as Temporal Pattern Recognition.	-	-	-	-	-	-	1-22
Licklider's Neural Autocorrelation Theory.	-	-	-	-	-	-	1-24
The Pitch Shift, Envelope or Fine-Structure.	-	-	-	-	-	-	1-28
The Existence Region of Residue Pitch.	-	-	-	-	-	-	1-30
Frequency Resolution and the Existence Region.	-	-	-	-	-	-	1-32
The Relative Importance of Low and High Harmonics:	-	-	-	-	-	-	-

Spectral Dominance.	-	-	-	-	-	1-34
The Effects of Relative Phase.	-	-	-	-	-	1-38
Relative Phase Effects and Auditory Nonlinearity.	-	-	-	-	-	1-42
The Pitch Shift and Auditory Nonlinearity.	-	-	-	-	-	1-45
The Residue Pitch of Dichotic Two-Component Complexes.	-	-	-	-	-	1-47
The Audibility of Residue Pitch and the Identifiability of Components.	-	-	-	-	-	-
	-	-	-	-	-	1-51
The Pitch of Modulated and Delayed Noise.	-	-	-	-	-	1-55
Summary.	-	-	-	-	-	1-58
Chapter 2. Some Theories of Residue Pitch Perception.						
Abstract.	-	-	-	-	-	2-1
Introduction.	-	-	-	-	-	2-3
Goldstein's Optimum-Processor Model.	-	-	-	-	-	2-4
Terhardt's Multiple Cue Theory.	-	-	-	-	-	2-8
Wightman's Autocorrelation Model.	-	-	-	-	-	2-9
Component Frequency Estimate Variability and Residue Pitch.	-	-	-	-	-	2-14
The Variability of an Estimated Residue Pitch at the Fundamental Frequency.	-	-	-	-	-	-
	-	-	-	-	-	2-15
Estimates of the Residue Pitch Difference Limen.	-	-	-	-	-	2-17
Computer Simulation Methods for Goldstein's and Wightman's Models.	-	-	-	-	-	2-22
Summary.	-	-	-	-	-	2-26

Chapter 3. A Decision Model for Pitch Discrimination, and Some Theories of Frequency Discrimination.

Abstract.	-	-	-	-	-	-	3-1
Introductory Comments.	-	-	-	-	-	-	3-3
The Signal Detection Model of Discrimination.	-	-	-	-	-	-	3-3
The Theoretical Psychometric Function.	-	-	-	-	-	-	3-4
The estimated Psychometric Function.	-	-	-	-	-	-	3-7
The Validity of the Statistical Assumptions.	-	-	-	-	-	-	3-8
The Application of the Decision Model to Component Frequency and Residue Pitch Discrimination.	-	-	-	-	-	-	3-9
What Factors are Likely to Affect Component Frequency Discrimination?	-	-	-	-	-	-	3-11
Frequency Discrimination for Single Sinusoids.	-	-	-	-	-	-	3-12
The Energy Detector Model.	-	-	-	-	-	-	3-14
Predictions of the Energy Detector Model.	-	-	-	-	-	-	3-18
Temporal Theories of Frequency Discrimination.	-	-	-	-	-	-	3-21
The Signal Detection Model and Temporal Theories of Frequency Discrimination.	-	-	-	-	-	-	3-31
Summary.	-	-	-	-	-	-	3-32

Chapter 4. Introduction to the Experiments, and Experiments 1 and 2.

Abstract.	-	-	-	-	-	-	4-1
The Choice of Stimulus Parameters.	-	-	-	-	-	-	4-3
The Experiments to be Described.	-	-	-	-	-	-	4-5
Apparatus for the Experiments.	-	-	-	-	-	-	4-7

General Properties of the Stimuli.	-	-	-	4-9
The Purpose of Experiments 1 and 2.	-	-	-	4-10
Method for Experiments 1 and 2.	-	-	-	4-11
Results of Experiments 1 and 2.	-	-	-	4-18
The Effect of Presentation Level on Discrimination for Residue	-			
Pitch.	-	-	-	4-20

Chapter 5. The Estimation of the Component Frequency Variability

Parameter $\sigma(f'_j)/f_j$.

Abstract.	-	-	-	-	-	5-1
Introductory Comments.	-	-	-	-	-	5-3
Problems of Experimental Method.	-	-	-	-	-	5-4
Experiment 3 - Method.	-	-	-	-	-	5-7
Results.	-	-	-	-	-	5-11
The Estimates of $\sigma(f'_j)/f_j$.	-	-	-	-	-	5-13

Chapter 6. The Estimation of the Residue Pitch Variability

Parameter $\sigma(f'_0)/f_0$.

Abstract.	-	-	-	-	-	6-1
Introductory Comments.	-	-	-	-	-	6-3
Comments on Method for Experiment 5.	-	-	-	-	-	6-4
Comments on Method for Experiment 6.	-	-	-	-	-	6-8
Experiments 4, 5 and 6 - Method.	-	-	-	-	-	6-11
Experiments 4, 5 and 6 - Results.	-	-	-	-	-	6-15
The Estimates of $\sigma(f'_0)/f_0$.	-	-	-	-	-	6-19
Comparison of $\sigma(f'_0)/f_0$ and $\sigma(f'_j)/f_j$.	-	-	-	-	-	6-20
The Discrepancy Between Estimates of $\sigma(f'_0)/f_0$ from Two	-					-

Different Tasks.	-	-	-	-	-	6-21
General Comments.	-	-	-	-	-	6-23
Summary.	-	-	-	-	-	6-24

Chapter 7. Component Frequency and Residue Pitch Variability
as a Function of Signal-to-Noise Ratio.

Abstract.	-	-	-	-	-	7-1
Introductory Comments.	-	-	-	-	-	7-5
Method for Experiments 7, 8, 9, 10 and 11.	-	-	-	-	-	7-8
Results of Experiments 7, 8, 9, 10 and 11.	-	-	-	-	-	7-15
Estimates of $\sigma(f'_j)/f'_j$.	-	-	-	-	-	7-19
Estimates of $\sigma(f'_0)/f'_0$ and the Theoretical Predictions.	-	-	-	-	-	7-19
The Effect of Frequency Uncertainty.	-	-	-	-	-	7-20
Evidence for an Internal Contribution to $\sigma(f'_0)/f'_0$.	-	-	-	-	-	7-21
The Importance of Between-Trial Comparisons.	-	-	-	-	-	7-23
The Dependence of Discrimination of Signal-to-Noise Ratio.	-	-	-	-	-	7-24
Summary.	-	-	-	-	-	7-25

Chapter 8. Residue Pitch Variability as a Function of Signal
Duration.

Abstract.	-	-	-	-	-	8-1
Introductory Comments.	-	-	-	-	-	8-3
Method for Experiments 12 and 13.	-	-	-	-	-	8-5
Results of Experiments 12 and 13.	-	-	-	-	-	8-8
A Comparison of the Effects of Duration and Signal-to-Noise Ratio.	-	-	-	-	-	8-11
The Differences between Discrimination Performance in Two	-	-	-	-	-	-

Tasks.	-	-	-	-	-	-	8-13
Summary.	-	-	-	-	-	-	8-14

Chapter 9. A Theoretical Discussion of the Discrimination Experiments.

Abstract.	-	-	-	-	-	-	9-1
The Predictions of Residue Pitch Theories and the Two	-	-	-	-	-	-	-
Fundamental Frequency Discrimination Tasks.	-	-	-	-	-	-	9-3
A Modification to Goldstein's Model.	-	-	-	-	-	-	9-9
Estimates of the Pitch Processor Noise.	-	-	-	-	-	-	9-11
The Fit of the Modified Model to the Discrimination Data.	-	-	-	-	-	-	9-13

Chapter 10. The Identification of Residue Pitch as a Function of Signal-to-Noise Ratio and Signal Duration.

Abstract.	-	-	-	-	-	-	10-1
Introduction.	-	-	-	-	-	-	10-3
The Derivation of Predictions.	-	-	-	-	-	-	10-6
The Matched Pitches of Two-Component Signals.	-	-	-	-	-	-	10-9
Method for Experiments 14 and 15.	-	-	-	-	-	-	10-9
Results of Experiments 14 and 15.	-	-	-	-	-	-	10-16
Analysis of Variance of Observed $\underline{d'}$.	-	-	-	-	-	-	10-17
Comparison of Observed and Predicted $\underline{d'}$.	-	-	-	-	-	-	10-23
The Observed and Predicted Distribution of Pitch Matches.	-	-	-	-	-	-	10-26
Summary.	-	-	-	-	-	-	10-27

Chapter 11. A General Discussion.

Introduction.	-	-	-	-	-	-	11-1
Discussion of the Experimental Results in terms of Theories of	-						
Pitch Perception.	-	-	-	-	-	-	11-4
Conclusions with respect the Theories of Pitch Perception.							11-6
Future Research on Theories of Pitch Perception.						-	11-7
The Experimental Results and Theories of Frequency						-	-
Discrimination.	-	-	-	-	-	-	11-9
Future Research on Theories of Frequency Discrimination.							11-10
Limits to the Generality of the Experimental Findings.							11-10
Other Applications of Spectral-Pattern Recognition.						-	11-11
Probabilistic Pattern Perception.				-	-	-	11-11

Bibliography.

Appendix A.

Glossary of Terms.

Chapter 1 reviews the perception of pitch, with particular reference to the 'residue' pitch of complex tonal stimuli. The review considers those physical properties of sound which may be logically related to pitch, and, from a selective discussion of the published literature, attempts to critically evaluate the importance of the possible physical correlates of pitch. The possibility that simple transformations of sounds caused by auditory non-linearities may confound the supposedly independent manipulation of certain physical parameters is also considered. The Chapter provides a general discussion of the 'periodicity' theory due originally to Schouten (1938, 1940), the 'neural autocorrelation' theory described by Licklider (1959), and a 'spectral-pattern recognition' theory exemplified by two formal models developed by Wightman (1973b) and by Goldstein (1973). The reviewed studies indicate that residue pitch is primarily determined by the frequencies of resolved components rather than by periodicity, and it is concluded that the spectral-pattern theories are able to account for the majority of the available data.

Chapter 2 discusses in detail the models proposed by Goldstein (1973) and Wightman (1973b). A general property of the spectral-pattern theories is that the residue pitch of a tonal complex can be considered as a random variable derived more or less directly from a weighted sum of a series of random variables representing internal estimates of the frequencies of

the components of the signal. In contrast, the residue pitch predicted by Schouten's periodicity theory and by an unmodified version of Licklider's neural autocorrelation theory can be considered to be derived from a single random variable, this being an internal estimate of the period of an unanalysed residue. This distinction, together with the assumption that the components of a complex signal are well-resolved, leads to a psychophysically testable distinction between the spectral-pattern and the periodicity theories. According to the spectral-pattern models, $\sigma(f'_0)/f_0$, the relative standard deviation of the theoretical distribution of the residue pitch in the region of the fundamental frequency of a harmonic signal, should be smaller than the minimum of $\sigma(f'_j)/f_j$, the relative standard deviations of the theoretical internal distributions of estimates of the component frequencies. The periodicity theory leads to the reverse prediction.

Predictions are also developed from the spectral-pattern models for the pitch probability density function (pdf) of the two two-component signals employed in the experiments described in the subsequent Chapters. If $\sigma(f'_j)/f_j$ is known for each component of these signals, a computer simulation of these models can predict the areas of each mode of the multimodal pitch pdf. Estimates of the parameter $\sigma(f'_j)/f_j$ from musical interval identification tasks and from frequency discrimination for complex harmonic signals are discussed.

Chapter 3 introduces the statistical method to be employed for the estimation of the parameters $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ from Two-Alternative-Forced-Choice (2AFC) pitch discrimination tasks. The method is based upon the Theory of Signal Detectability (Green and Swets, 1966). This chapter also reviews some theories of frequency discrimination for simple sinusoidal signals, including an Energy Detector model (Green and Swets, *ibid*; Henning 1967a) based upon spectral analysis, a temporally based model due to Siebert (1970), and modifications of that model described by Luce and Green (1974) and Goldstein and Srulovicz (1977). These theories are generalized to the discrimination of a residue pitch derived from several well-resolved components of a complex signal.

Chapter 4 describes the stimuli and apparatus employed throughout the experimental studies. This Chapter also describes two preliminary studies. Experiment 1 is an investigation of the effects of presentation level upon discrimination between the residue pitches of two signals composed of the 4th and 5th, and the 6th and 7th harmonics of 200 Hz. Experiment 2 is a study of frequency discrimination for a single component at 1200 Hz. The results of Experiment 1 indicate that two of the three observers had some difficulty ignoring the frequency differences between the components of the two signals. The effects of presentation level on discrimination were generally comparable in both tasks, and also comparable

with the results described by Weir, Jesteadt, and Green (1977).

The discussion of the pitch discrimination experiments in the thesis is ordered differently from the discussion immediately below. Chapters 5 and 6 contain descriptions of Experiment 3 to 6, which investigated the effects of signal level on component and fundamental frequency discrimination. Chapter 7 describes Experiments 7 to 11, where signal-to-noise ratio was varied, and Chapter 12 describes Experiments 12 and 13, where signal duration was varied.

The first stage of the main experimental investigation consisted of the estimation of the parameter $\sigma(f'_j)/f_j$. Chapter 5 describes Experiment 3, a 2AFC frequency discrimination experiment intended to estimate the parameter $\sigma(f'_j)/f_j$ as a function of sensation level for both components of a signal containing the 4th and 5th harmonics of 200 Hz. One component was presented as a tone burst, while the second was continuously present, as was a low-pass filtered noise intended to mask any auditory distortion products at frequencies below the signal components. The absolute frequency difference between the two tone-bursts in each trial was constant for each block of 50 trials. Two observers took part in this study.

The obtained estimates were not a simple monotonically decreasing function of presentation level, but decreased with

increasing level for signals up to 30 dB SL, and showed a marked increase at 50 dB SL. This nonmonotonicity was attributed to two factors. Firstly, for signal levels greater than 30 dB SL, the masking noise led to a constant signal-to-noise ratio, which would be expected to give rise to a constant DL. The increase in the DL at 50 dB SL was assumed to arise from an increased upward spread of masking from the low-pass filtered noise. The resulting estimates of $\sigma(f'_j)/f_j$ ranged from about 0.002 to 0.008.

A similar experiment, Experiment 7, which is described in Chapter 7, estimated $\sigma(f'_j)/f_j$ as a function of signal-to-noise ratio for a single observer. The level of the signal components was constant, and the spectral-density of a white noise was varied between 20 and 40 dB below the level of the components. The obtained estimates of $\sigma(f'_j)/f_j$ were monotonically dependent on the signal-to-noise ratio, with a slope of approximately $1/\sqrt{N_o}$, and are consistent with an Energy Detection model of frequency discrimination. The estimates ranged from about 0.012 at a signal-to-noise ratio of 20 dB to about 0.003 at 40 dB.

The estimates of $\sigma(f'_j)/f_j$ from Experiments 3 and 7 are only slightly larger than the relative frequency-difference limen for single sinusoids, and were typically smaller than values estimated by other workers from musical interval

identification.

The second stage of the investigation consisted of a number of experiments designed to estimate the relative fundamental frequency-difference limen $\sigma(f'_0)/f_0$, under conditions as similar as possible to those pertaining to the estimates of $\sigma(f'_j)/f_j$.

Two different experimental methods were employed for the estimation of $\sigma(f'_0)/f_0$ for f_0 at approximately 200 Hz. Each method employed a 2AFC pitch discrimination task, where the absolute fundamental frequency difference $\{\Delta f\}$ was constant during each block of trials. In Method One, observers were asked to compare the residue pitches of two two-component signals containing the same two harmonics of slightly differing fundamental frequencies. This method has the advantage of minimizing the effects of frequency uncertainty (Green and Swets, 1966), but does not ensure that correct performance is due to the use of a difference in the residue pitches of the two signals, since the pair-wise between-signal component frequency differences are completely correlated with the fundamental frequency difference.

Method Two employed two signals of differing harmonic structure. The observers compared the pitch of a reference signal containing the fundamental component and the 2nd, 7th,

8th, and 9th harmonics with that of a two-component comparison signal containing the 4th and 5th or the 5th and 6th harmonics of a slightly differing fundamental frequency. The signal containing the 5th and 6th harmonics was presented to discourage the observers from using cues based on between-trial component frequency differences. Here, correct performance does ensure that the observer was basing his judgement on the residue pitch of the two-component signal, but the larger range of frequencies covered by the components of the two signals may lead to a degree of frequency uncertainty. Estimation of $\sigma(f'_0)/f_0$ for the residue pitch of the comparison signal from this method requires that the same parameter is known for the reference signal. Experiment 4, which is described in Chapter 6, fulfilled this function.

Experiment 6, which is described in Chapter 6, estimated $\sigma(f'_0)/f_0$ as a function of the sensation level of the components for two observers using Method One. The obtained estimates showed the same nonmonotonic dependence on sensation level as the estimates of $\sigma(f'_j)/f_j$ from Experiment 3, and were typically smaller than the smaller estimate of $\sigma(f'_j)/f_j$ for comparable presentation conditions. This result is thus consistent with the spectral-pattern models and with the estimates of $\sigma(f'_j)/f_j$ from Experiment 3, and contrary to the prediction of the periodicity theory. The obtained result also demonstrates that the observers were employing information from

both components of the signal, and not simply basing their responses on the frequency difference at one of the two components. Experiment 10, which is described in Chapter 7, estimated $\sigma(f'_0)/f_0$ by the same method, as a function of signal-to-noise ratio, again with two observers. A comparison of the data from this experiment, for the one observer who took part in Experiment 7, with the estimates of $\sigma(f'_j)/f_j$ from the latter experiment, shows that the estimates of $\sigma(f'_0)/f_0$ from Experiment 10 are quantitatively consistent with the prediction from the spectral-pattern models and the estimates of $\sigma(f'_j)/f_j$ under comparable presentation conditions. This result confirms the finding from a comparison of the results of Experiments 3 and 6. The estimates of $\sigma(f'_0)/f_0$ from Experiment 10 are consistent with the prediction of a multiple-band Energy Detector employing two filters centered at about 800 and 100 Hz, whose equivalent rectangular bandwidths are between 11% and 15% of their centre frequency. These bandwidths coincide with estimates of the critical bandwidth at the same frequencies (Scharf, 1970).

Experiment 5, described in Chapter 6, employed Method Two for the estimation of $\sigma(f'_0)/f_0$ as a function of the sensation level of the components of the two-component signal. The results for the two observers showed a comparable nonmonotonic dependence on sensation level to that found in Experiments 3 and 6, but the estimates of $\sigma(f'_0)/f_0$ for the comparison containing

the 4th and 5th harmonics exceeded those from Experiment 6 by a factor of 1.51 for one of the two observers, and 1.71 for the second. These estimates were also larger than the smaller estimate of $\sigma(f'_j)/f_j$ from Experiment 3 at the same SL. An allowance for the stimulus uncertainty expected from Method Two reduced, but did not eliminate the discrepancy between the results of Experiments 5 and 6. The values of $\sigma(f'_j)/f_j$ required to account for the estimates of $\sigma(f'_0)/f_0$ from Experiment 5 are up to twice the values estimated in Experiment 3.

Experiment 8, described in Chapter 7, estimated $\sigma(f'_0)/f_0$ by Method Two as a function of signal-to-noise ratio. The estimates for the comparison signal containing the 4th and 5th harmonics in this experiment exceeded those from Experiment 10, where Method One was employed, by a factor of 1.38 for one of the observers, and 2.13 for the other. The dependence on signal-to-noise ratio was comparable for the two sets of estimates, for both observers and for signal-to-noise ratios between 20 and 30 db.

Some part of the discrepancy between the estimates of $\sigma(f'_0)/f_0$ from the two methods employed here may simply be due to the effect of frequency uncertainty. Experiment 9 was performed to examine this possibility. This experiment differed from Experiment 8 in that only the two-component comparison

signal containing the 4th and 5th harmonics occurred, so that frequency uncertainty for this signal was eliminated. Only one observer was available for this experiment. The estimates of $\sigma(f'_0)/f_0$ obtained from Experiment 9 were smaller than those from Experiment 8 by an average factor of about 1.5, which is slightly larger than the factor of 1.22 expected from the uncertainty difference, but confirms the assumed uncertainty effect. The estimates of $\sigma(f'_0)/f_0$ from Experiment 9 are approximately equal to the larger of the estimates of $\sigma(f'_j)/f_j$ for the same observer from Experiment 7, and exceed the predictions of the spectral-pattern models based on the latter estimates.

The contribution of cues from between-trial differences in the component frequencies of the comparison signals in Method Two was assessed by Experiment 11, which is described in Chapter 8. Here, a number of different values of $|\Delta f|$ occurred randomly within a block of trials; the experiment was otherwise similar to Experiment 8. A comparison of the results of Experiments 8 and 11 shows that between-trial comparisons are not an important determinant of the discrimination performance measured by Method Two. The random presentation of different values of $|\Delta f|$ within a block of trials had only slight effects upon the observed performance, consistent with an increased variability of the decision criterion, and there was no significant difference in the dependence of discrimination of signal-to-noise ratio.

Chapter 8 contains a description of two experiments intended to estimate the effects of signal duration on $\sigma(f'_0)/f_0$. Experiment 12 employed Method Two, and Experiment 13 employed Method One. The components of the two-component harmonic signal were presented at about 30 dB SL, in a background of noise whose spectral-density was 30 dB below the amplitude of those components. The duration of the two-component signal was varied between 20 and 80 msec.

The observed duration dependence of $\sigma(f'_0)/f_0$ was consistent with an inverse dependence upon the square root of duration for durations between 25 and 80 msec, while for a duration of 20 msec, $\hat{\sigma}(f'_0)/f_0$ increased to a value greater than that predicted by that relationship. The differences between the estimates of $\sigma(f'_0)/f_0$ from the two experiments were comparable with the difference between Methods One and Two described above.

The estimates of $\sigma(f'_0)/f_0$ from Experiments 12 and 13 were compared with those obtained from Experiments 8 and 10, where signal amplitude and duration were constant while noise power was varied. At equivalent signal-to-noise energy ratios, the estimates of $\sigma(f'_0)/f_0$ were somewhat larger for the shorter duration signals. These differences are discussed in terms of differences between the rise and fall times of the signals, and in terms of a modified Energy Detector model with an additive

internal variance.

Chapter 9 reviews the results of the pitch and frequency discrimination experiments described in the previous four chapter primarily in terms of Goldstein's (1973) spectral-pattern recognition theory of residue pitch. The differences between estimates of $\sigma(f'_0)/f_0$ from Methods One and Two are interpreted in terms of Goldstein's assumption that the representation of component frequency at the pitch processor includes an internal noise component. Such a noise component could represent either, as Goldstein has proposed, a contribution from noisy frequency-selective channels, or a noisy criterion for comparison of pitches derived from non-coincident harmonics. Estimates of the extent of such an internal noise are made from the pitch discrimination data. It is concluded that pitch comparisons which may be made by reference to multiple within-channel comparisons of frequency differences, such as those for the experiments employing Method One, are unaffected by such an internal noise, and that the variability of such judgements is due only to the variability of the peripheral coding of frequency and to external noise.

However, where pitch judgements are made between two signals affecting two different sets of frequency selective channels, as in the experiments employing Method Two, the postulated internal noise would lead to a performance decrement,

in agreement with the results described above. A modification of Goldstein's model is described, allowing the sign of a pitch difference to be estimated from within-channel frequency differences before the pitch itself is extracted by the central processor. This modification allows the model to provide an excellent account of the performance of one observer, and a fair account of the performance of a second. A comparable modification to Wightman's (1973b) autocorrelation model would have the same consequences.

The value of the postulated internal noise was estimated from performance in experiments requiring both within and between-channel comparisons. On a scale of relative frequency, the internal noise variance was estimated as about 10^{-5} . The corresponding standard deviation is about three times smaller than the value of 0.01 assumed by Goldstein (1973) in describing performance in musical interval identification. Goldstein's estimation of this parameter would, however, include any external noise, the noise due to peripheral coding, and also an effect of frequency uncertainty.

Chapter 10 contains a description of two experiments intended to estimate the effects of both signal-to-noise ratio and signal duration, through their effect on $\sigma(f'_j)/f_j$, upon the relative areas of the major modes of the internal representation of the residue pitch of a two-component harmonic

signal. These experiments, Experiments 14 and 15, were essentially similar to each other, and employed the same signals and stimulus conditions as Experiments 8 and 12. The two observers made same-different judgements for the pitches of a standard and a comparison signal. The standard signal contained the 4th and 5th, or the 5th and 6th harmonics of 170, 205 or 258 Hz. The comparison signal contained components at its fundamental frequency, and also at the 2nd, 7th, 8th, and 9th harmonics. The fundamental frequency of the comparison signal was equal to the fundamental frequency of the standard signal or to a frequency corresponding to the mean of one of the theoretical pitch modes above and below the standard fundamental frequency. The pitch modes above and below the fundamental frequency were derived from Goldstein's (1973) model; these pitches do not differ significantly from those predicted by Wightman's (1973b) model.

The obtained results were compared with the predictions derived from computer simulations of Goldstein's and Wightman's models. The simulations were performed with estimates of $\sigma(f'_j)/f_j$ based upon the results of Experiments 8 and 12. The estimates used for the simulation included an internal noise contribution, and an allowance for the frequency uncertainty occasioned by the stimulus materials for Experiments 14 and 15. The observed and predicted data were compared both in terms of the probability of an acceptance of a given pitch match, and in

terms of the sensitivity index d' . Although the results were subject to the effects of practice, and to the observer's familiarity with the stimulus material, performance was generally consistent with the predictions of both models.

Chapter 11 summarizes the results of the experiments, and considers their implications for theories of pitch perception and theories of frequency frequency discrimination. A number of possibilities for future research in each of these areas are outlined. The chapter concludes with a brief general discussion of the possible applications of similar probabilistic pattern recognition theories for other stimuli.

CHAPTER 1

Abstract.

This chapter is intended to provide a review of the perception of pitch, with particular reference to the 'residue' pitch of complex sounds. The phenomenon of residue pitch has also been described by the terms 'the missing fundamental', 'fundamental tracking', 'low pitch' and 'virtual pitch'. The particular auditory stimuli considered here are those containing periodic frequency components; pitch-like sensations arising from stimuli constructed from aperiodic noise have also been reported, and such pitch sensations are discussed briefly. The chapter begins with a discussion of those physical parameters of sound which may be logically related to pitch, and subsequently, a selective review of the published literature is presented. This review attempts to critically evaluate the importance of the possible physical correlates of pitch, and also considers the possibility that simple transformations of sounds caused by auditory non-linearities may have the effect of confounding the supposedly independent manipulation of certain physical parameters. Several theories of pitch perception are discussed in general terms, notably the 'periodicity' theory due originally to Schouten (1940), the 'neural autocorrelation'

theory described by Licklider (1959), and a 'spectral-pattern recognition theory' which is exemplified by two formal models developed by Wightman (1973b) and by Goldstein (1973).

Introductory Remarks.

Auditory events provide a rich source of information, and our ability to identify sounds, and to differentiate between them, is necessarily fundamental to our perception of speech, music, and all other audible events. Pitch is one of the primary perceptual attributes of sound, and the variation of pitch between certain points along a scale of subjective height, that is, those points having a representation on some musical scale, is essential for the production of melody. The development of an adequate psychological theory of pitch perception, that is, a theory which can generate reasonably accurate predictions concerning the perceived pitch of any acoustic signal, requires a knowledge of the relations between the physical parameters of sounds and the attribute of pitch. Such a relationship may be simple and direct, or complex and also subject to influence by other non-acoustic factors; in either case it will be necessary to define a set of rules governing the extraction of pitch from the physical stimulus.

Musicians and musical instrument makers have known for very many years that the pitch produced by a vibrating stretched string or column of air becomes higher as the length of the string or air column is decreased; it is also well known that a change in the length of such a vibrating element produces an inversely proportional change in the rate at which the oscillation of the elements repeats over time. This change in repetition rate appears to be closely related to changes in pitch.

Repetitive or 'periodic' oscillations may be physically described in terms of two alternative and generally equivalent representations in the related domains of time and frequency. These two representations are discussed in detail elsewhere (Littler, 1965), but a brief and approximate description is given here. The waveform, the temporal representation, simply indicates the form of the change over time of the instantaneous amplitude of the oscillation of the signal. The waveforms shown in figure 1.1a) are approximations to the waveform produced by the bowing of a violin string, vibrating at frequencies of 200 and 300 Hz. Three salient features of the waveform are apparent in the figure. The 'period' t_0 is the time interval over which the pattern of amplitude change is repeated. The 'envelope', is the gross repetitive pattern formed by the instantaneous amplitude peaks, which in the case of this signal repeats in the same time interval t_0 as the waveform itself. The 'fine-structure' consists of the details of the pattern of amplitude changes over time intervals less than the period. The frequency domain representations, or spectra of the same signals are shown in figure 1.1b); each vertical line indicates the presence of energy at a particular frequency. This alternative representation of the signal is possible because of the principles of Fourier analysis by which it is known that any periodic oscillation may be decomposed into a series of simple or sinusoidal oscillations at frequencies which are integer multiples of the repetition rate or 'fundamental' frequency of the signal. The spectra in figure 1.1b) each show such a series of simple components. This simple description neglects the phase spectrum of the signal, but the subsequent discussion will

Figure 1.1 a). Schematic representations of the waveform produced by a violin producing pitches corresponding to 200 Hz (i) and 300 Hz (ii).

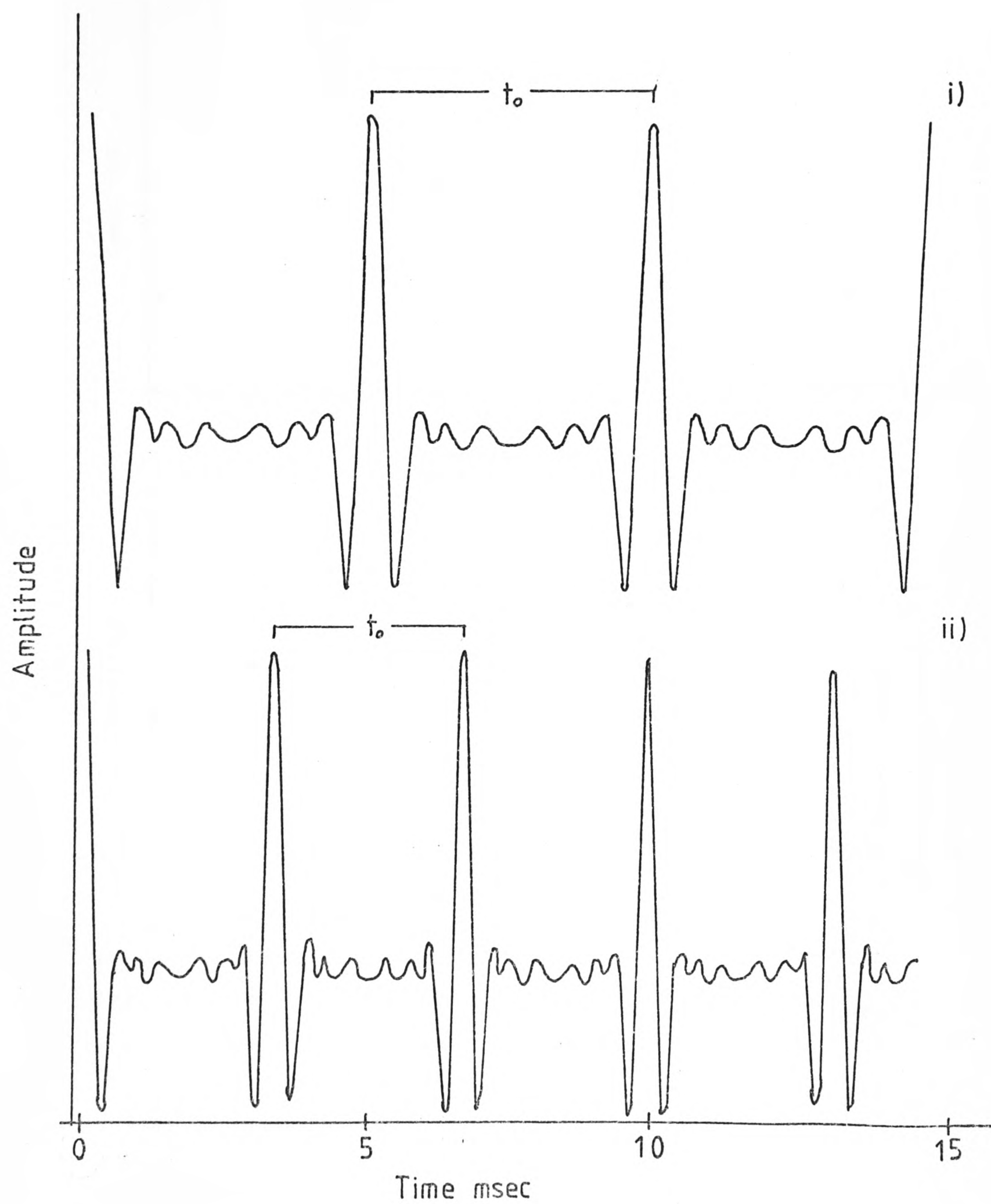


Fig. 1.1a)

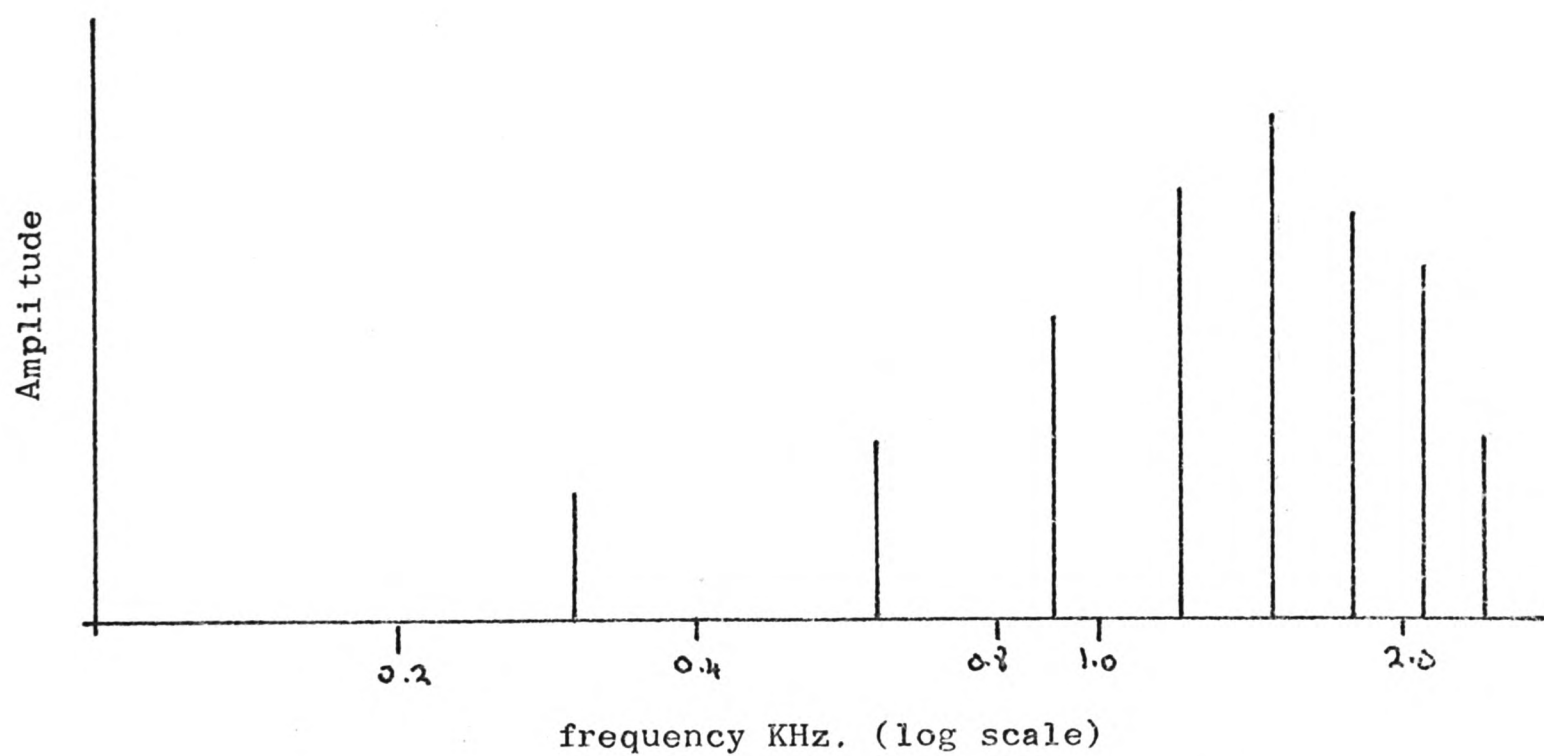
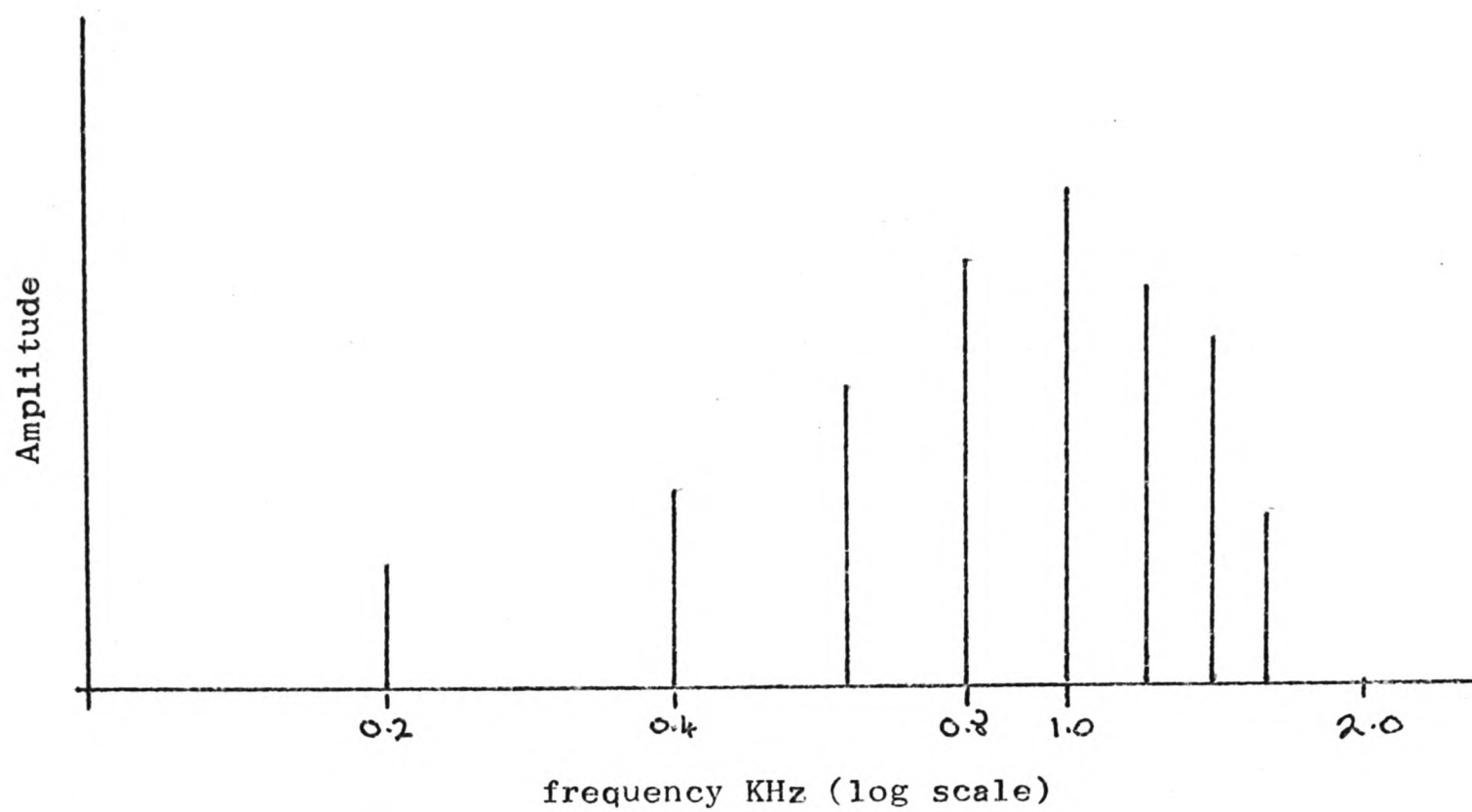


Figure 1.1b). Amplitude spectra of the two signals shown in figure 1.1a).

take phase into account.

The increase in pitch resulting from the shortening of a vibrating violin string corresponds to a number of changes in both temporal and spectral parameters of the sound. The period, the envelope, and time intervals in the fine-structure of the waveform all decrease in proportion to the decrease in length, while the fundamental frequency, the values of the individual component frequencies, and the differences between those frequencies all increase in inverse proportion to the decrease in length. An acoustic signal such as this is thus rather uninformative in that the parameters which may be logically related to pitch are all correlated, and no isolation of any particular parameter is possible. Fortunately, artificial signals may be constructed in such a way that some of these parameters become relatively independent, and such signals are much used in the experimental work discussed below.

The Quantification of Pitch.

The relation between perceived pitch and some physical parameter of the acoustic signal may be better described if some means is available for quantifying pitch. For a simple sinusoidal signal, all of the above temporal and spectral parameters are completely correlated, so that the pitch of a sinusoid may be arbitrarily assigned the value of the frequency in Hz of the signal. In order to quantify the pitch of some non-sinusoidal signal, the frequency of a sinusoid may be adjusted until the two signals are of equal pitch, and the pitch

of the non-sinusoidal signal is then assigned a pitch value corresponding to the frequency of the sinusoid, and also scaled in Hz.

Psychological Theories of Pitch Perception.

A psychological theory of pitch perception may be conveniently thought of as consisting of two portions. The first part of such a theory will be the assumption that one or more of the above-mentioned physical parameters is effective in the determination of pitch, while the second part will contain a formulation of the supposed rules governing the derivation of a particular pitch value from the sensory information provided by the effective physical parameter or parameters. The first part of such a theory may be rejected if it can be shown that the chosen parameter, or combination of parameters, is not, under all conditions, related to pitch, or if it can be demonstrated that the nature of auditory physiology is such that the auditory system is insensitive to the chosen parameter or parameters under some conditions which do not prohibit the perception of pitch. The second part of such a theory may be rejected if it can be shown that perceived pitch does not correspond to predictions made in accordance with the assumed rules.

Unfortunately, the interpretation of both behavioural and physiological data presents a number of difficulties. Physiological observations are made upon the auditory systems of a variety of species, but very rarely upon humans, and there are almost certainly significant inter-species differences in the

structure and organization of the auditory system (Whitfield, 1978). Furthermore, the complexity of neurophysiological systems is such that the existing evidence is not easily related to theories of human perception, particularly since these theories are intended to explain behavioural data. For these reasons, it may not be reasonable to conclude from physiological data that some particular information cannot be coded, or that the accuracy of the neural representation of some particular information in humans is well described by estimates based upon data from animals. Nevertheless, if the auditory physiology of a range of species shows some general property, then it is reasonable to assume that this same property would be evident in the human species. Thus, evidence from physiological and neurophysiological studies can serve as support for the plausibility of some presumed process, but the absence of support for some presumed process from such studies cannot be taken as demonstrating that no such process occurs.

One major difficulty in the interpretation of behavioral evidence arises from the unknown and complex relationship between an observer's perception and his overt response; this problem is compounded here by the subjective variability of the pitch percept. All observers possessing normal auditory acuity are likely to agree on the pitch produced by a musical instrument, and if these observers are able to control the pitch of their voices, they are also likely to be able to reproduce that pitch by singing; however, the sounds constructed in the laboratory for the purposes of testing theories of pitch perception tend, probably because the pitch cues are fewer or

weaker, to be perceived differently by different observers. Because of this, behavioural evidence is generally based upon judgements made by selected observers who have, through training and practice, acquired the ability to respond with reasonable consistency; nevertheless, observers may disagree, and consequently, it is unwise to place too great a weight upon evidence based on the performance of a small sample of observers.

The experimental investigations reported below are interpreted in terms of a statistical model of the internal representation of the pitch of a stimulus. The basis for this model is introduced in Chapter 3.

An Early Theory, Ohm's Acoustical Law.

In 1843, some twenty years after the mathematical basis of spectral analysis had been formulated by Fourier, G.S. Ohm proposed a hypothesis concerning the relationship between perceived pitch and the frequencies of the spectral components of an acoustic signal. This hypothesis, often called Ohm's Acoustical Law, stated:-

- "a) To produce a tone of frequency \underline{m} , a series of periodic impressions with a period equal to $1/\underline{m}$ is necessary; the form $\underline{a \cdot \sin 2\pi(mt+p)}$ must be present in each period either purely or as a real component.
- b) These forms must have the same value of \underline{p} in successive periods.

c) The value of a must be the same in successive periods." (as quoted in translation by de Boer (1976, p492)). The term frequency should be taken to mean pitch as defined above; the algebraic terms m, representing the frequency of some simple component, and p, representing the phase angle of that component, are elsewhere written as f and θ respectively; the term a is the amplitude of the component, and t represents time. The essence of Ohm's hypothesis was then, that a pitch corresponding to f Hz will be heard if and only if a component of frequency f Hz is present in the spectrum of the acoustic signal.

Helmholtz was much concerned with the ability of the auditory system to carry out an analysis of a complex sound into its simple spectral components. In "Die Lehre von den Tonempfindungen" (Helmholtz, 1863), he related his own observations of the audibility of the simple components of complex sounds, and claimed to be able, under favourable conditions, to hear independently each component up to the 16th harmonic, each being identifiable by its individual pitch. Helmholtz supposed that his inability to identify still higher components was due to their being so close in pitch that the pitches of these components could not be differentiated from each other. Thus, Helmholtz claimed that the series of simple components of a complex tone was audible as a series of pitches corresponding to the component frequencies.

The Place Theory, a Physiological Basis for Ohm's Law.

The conception of pitch perception propounded by Ohm and Helmholtz was thus essentially analogous to Fourier analysis, and implied that the pitches of simple sinusoidal components were the basic perceptual components from which more complex perceptions were compounded. Helmholtz also proposed a physiological mechanism capable of extracting such perceptual units from a complex sound, and representing the presence and intensity of these units as spatial distribution of mechanical activity. The general character of this mechanism was shown by Bekesy (1960) to be valid, although the details of Helmholtz's physiological model were wrong. Bekesy's extensive observations of the behaviour of the basilar membrane demonstrated that this membrane is frequency selective, and when set in motion by a complex oscillation, shows displacement peaks at points along its length which correspond uniquely with the frequencies of the Fourier components of the oscillation. The work of Helmholtz and Bekesy thus suggested a physiological analogue to Ohm's acoustic law, the 'place' theory, according to which the perception of a pitch corresponding to that of a sinusoid at the frequency f Hz requires a displacement of the basilar membrane at the place representing that frequency.

Seebeck, Schouten, and the Pitch of the Residue.

Evidence contrary to Ohm's Acoustical Law and the place theory was first presented by Seebeck in 1841 (Schouten, 1970; de Boer, 1976). By means of an acoustical siren, Seebeck was

able to generate periodic signals in such a way that the amplitude of each spectral component could be independently controlled. Seebeck noted that the subjective loudness of the pitch corresponding to the component at the fundamental frequency of a complex signal appeared to be independent of the physical amplitude of that component, and that this pitch was still audible when the amplitude of the fundamental component was almost zero. If this observation is correct, then it is clear that a component at the fundamental frequency is not a necessary condition for the perception of a pitch corresponding to that frequency, and unless a basilar membrane displacement at the corresponding place is somehow introduced in the absence of the appropriate Fourier component, the place theory must also be rejected. The precision of Seebeck's apparatus and techniques was, of necessity limited, so that doubts as to the reliability of this finding remained until Schouten (1938, 1940) was able to replicate Seebeck's experiment.

Schouten employed an optical siren to generate two signals of equal repetition rate, this being 200 Hz, one having the waveform of a pulse-train, and the other being a sinusoid. The sinusoid was equal in frequency and amplitude to the fundamental component of the pulse train, but of the opposite polarity, so that the sum of the two signals contained virtually no energy at the fundamental frequency. The waveforms and spectra of these two signals, together with those of the sum signal, are illustrated in figure 1.2 a), b) and c). Schouten showed that all three signals were perceived as being equal in pitch, despite the absence of the fundamental component in the sum

signal. This replication of Seebeck's result confirmed the insufficiency of Ohm's view of pitch perception, and led Schouten to develop an alternative theoretical framework. Schouten gave the name 'residue pitch' to this phenomenon of the pitch of a complex signal corresponding to the frequency of the missing fundamental component. Other terms for the same phenomenon have been used by other workers, notably 'periodicity pitch', 'the missing fundamental', 'fundamental tracking', 'virtual pitch' and 'low pitch'. Schouten's original term carries certain theoretical connotations which will become apparent shortly, but this term will be retained throughout this discussion since it remains the most commonly used of the alternatives.

Residue pitch is often considered to be a phenomenon which occurs only with the artificial signals produced in the laboratory, and while such phenomena are often of theoretical interest, they may be of limited relevance to a general understanding of the operation of the auditory system. However, an examination of the spectra and pitches of some of the sounds which are common outside the laboratory reveals that residue pitch is also part of everyday aural experience. Telephone systems and portable sound reproducers are generally unable to reproduce the lowest spectral components of speech and musical instruments in the lower registers; the pitches of these sounds so reproduced are nevertheless independent of the bandwidth limitations of the reproducing system. Also, several musical instruments, notably the bassoon, naturally generate sounds with little or no energy at the fundamental frequency, yet the

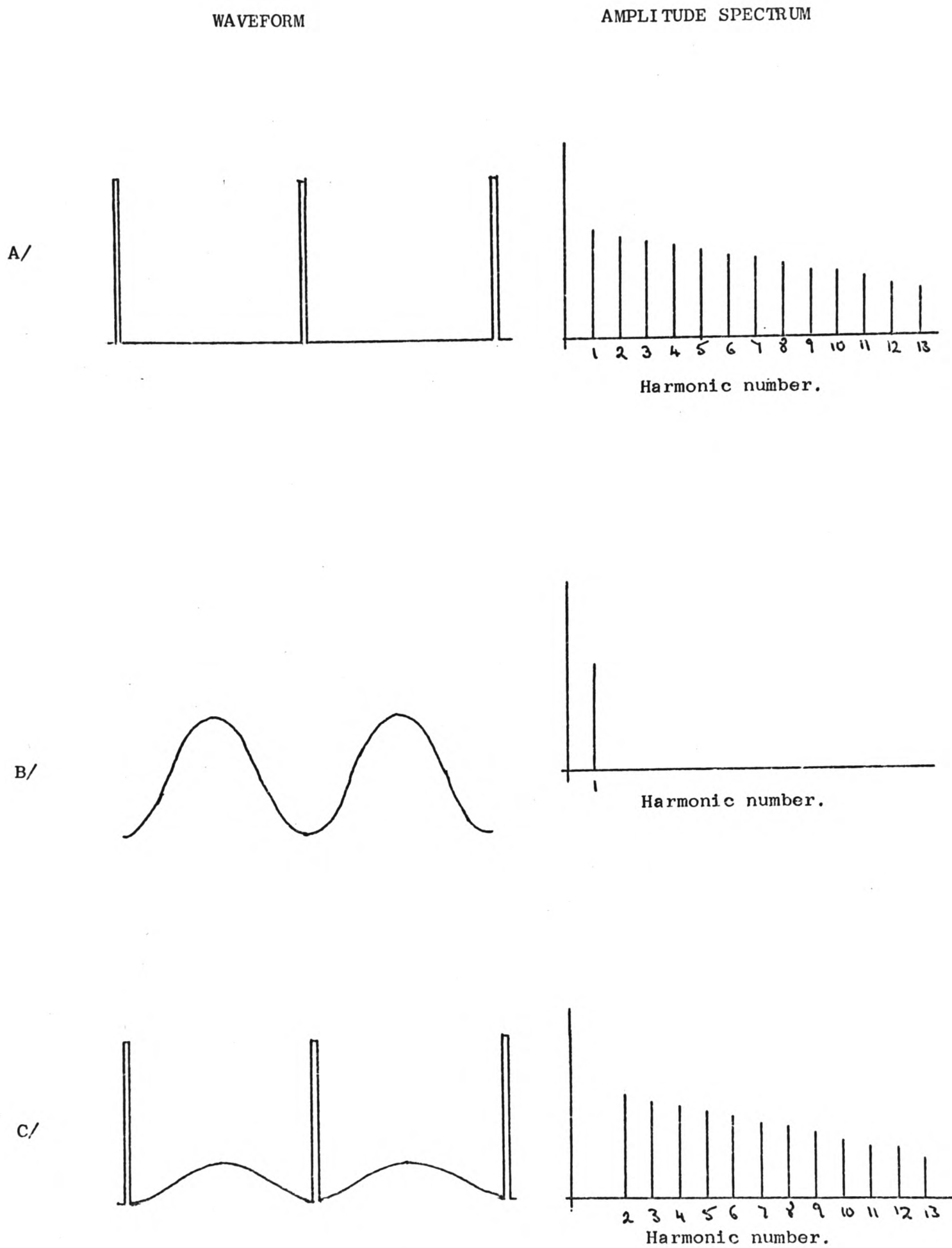


Figure 1.2. Schematic representations of the waveforms and spectra of the signals used by Schouten (1938). A/, periodic pulse train, B/ Sinusoidal component in opposite phase to component of equal frequency in A/. C/ Result of addition of A/ and B/.

signal. This replication of Seebeck's result confirmed the insufficiency of Ohm's view of pitch perception, and led Schouten to develop an alternative theoretical framework. Schouten gave the name 'residue pitch' to this phenomenon of the pitch of a complex signal corresponding to the frequency of the missing fundamental component. Other terms for the same phenomenon have been used by other workers, notably 'periodicity pitch', 'the missing fundamental', 'fundamental tracking', 'virtual pitch' and 'low pitch'. Schouten's original term carries certain theoretical connotations which will become apparent shortly, but this term will be retained throughout this discussion since it remains the most commonly used of the alternatives.

Residue pitch is often considered to be a phenomenon which occurs only with the artificial signals produced in the laboratory, and while such phenomena are often of theoretical interest, they may be of limited relevance to a general understanding of the operation of the auditory system. However, an examination of the spectra and pitches of some of the sounds which are common outside the laboratory reveals that residue pitch is also part of everyday aural experience. Telephone systems and portable sound reproducers are generally unable to reproduce the lowest spectral components of speech and musical instruments in the lower registers; the pitches of these sounds so reproduced are nevertheless independent of the bandwidth limitations of the reproducing system. Also, several musical instruments, notably the bassoon, naturally generate sounds with little or no energy at the fundamental frequency, yet the

pitches of these sounds are generally accepted as corresponding to the pitch of the missing fundamental component.

The above observations imply that residue pitch is the subjective occurrence of a pitch corresponding to the frequency of the absent fundamental component of a complex signal, but a more general definition is preferable, since not all complex signals are periodic, and thus not all complex signals have a true fundamental frequency, while they may still exhibit a pitch which corresponds to the frequency of some absent component. A residue pitch is defined here as a pitch produced by the joint action of two or more simple components, and which corresponds to some frequency lower than that of any of these components.

The Response of the Place Theorists: Auditory
Non-linearity.

The earliest hypothesis intended to reconcile Ohms' law and the place theory to the phenomenon of residue pitch was that the fundamental component was re-introduced into the ear by some non-linear distortion of the signal. One form of non-linearity which would re-introduce the fundamental component into a periodic complex signal is the quadratic non-linearity, where a signal $f(t)$ is transformed into

$$f'(t) = a_1 \cdot f(t) + a_2 \cdot f(t)^2$$

a_1 and a_2 being constants whose relative values depend upon the severity of the non-linearity. The quadratic non-linearity introduces distortion products at frequencies equal to the sums and differences between the frequencies of the original

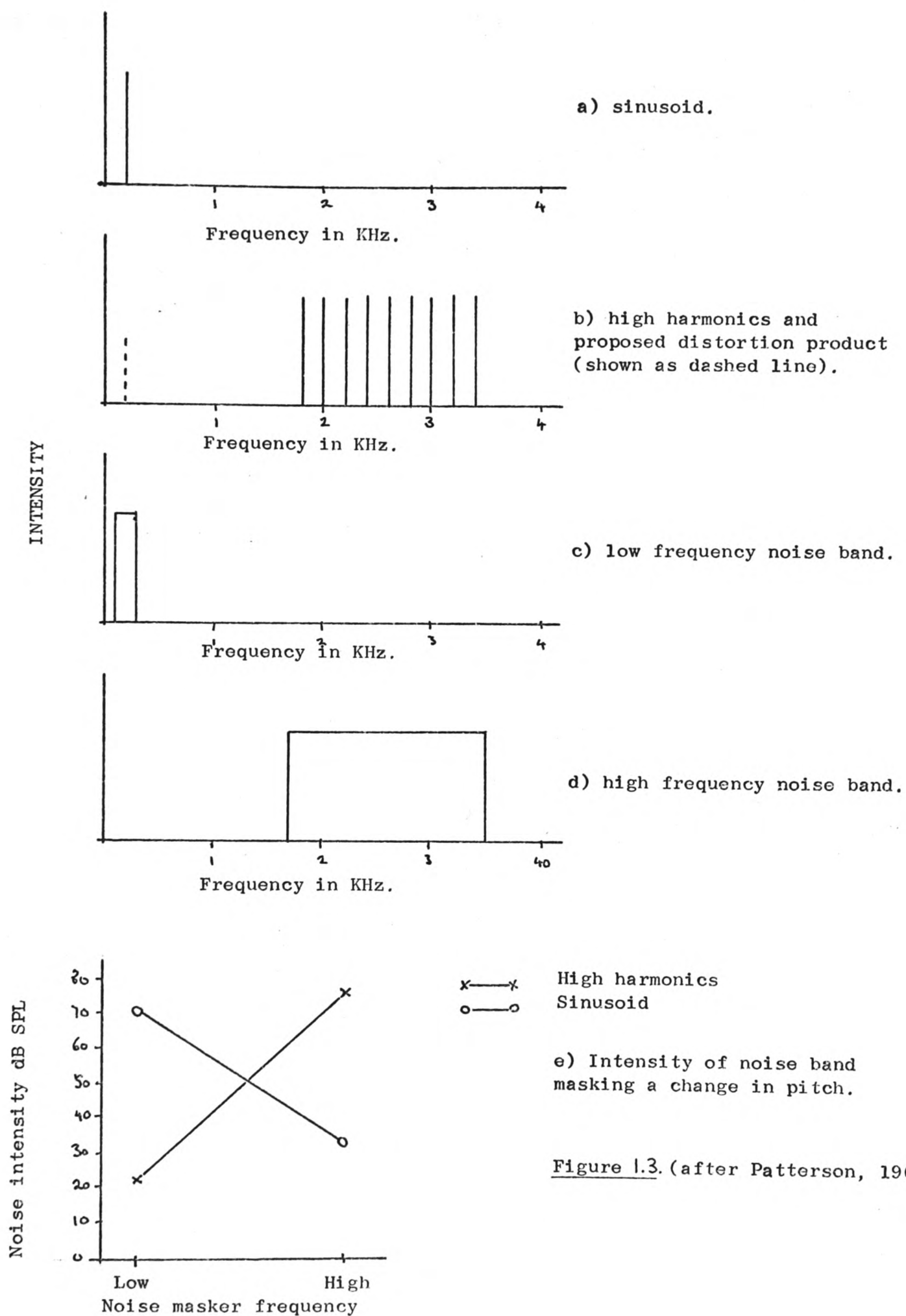
components of the signal, so that a periodic signal containing a series of harmonics , each differing in frequency from its neighbour by a value equal to the fundamental frequency, will yield a quadratic difference product at the fundamental frequency. This hypothesis, first suggested by Helmholtz as an explanation of Seebeck's observations of the apparent independence of the loudness of the pitch at the fundamental frequency and the amplitude of the corresponding spectral component, may now be rejected on the following grounds.

Schouten (1938) observed that the pitch produced by the sum signal described above (pp11-12), which corresponded to the pitch of a 200 Hz sinusoid, did not, when heard together with a sinusoid of 206 Hz, produce any audible beats. If a distortion product at 200 Hz had been acoustically present in the ear, then such beats would be expected. Furthermore, if a quadratic difference tone at the fundamental frequency were responsible for residue pitch, then this pitch should be masked by a band of noise which masks a sinusoidal signal of the same frequency and comparable loudness. Licklider (1954), Small and Campbell (1961), and Patterson (1969) have convincingly demonstrated that this is not the case. The task used by Patterson was the detection of a change in pitch from 240 to 280 Hz, this pitch being produced either by a single sinusoid, or by a pulse train, which was high-pass filtered so that the first ten harmonics (including therefore the fundamental component) were removed. Patterson presented each of these signals against a background of noise, the noise being band-pass filtered so that either frequencies close to the fundamental or around the high

harmonics were present, and measured the noise power required to mask the change in pitch of each signal. The results of this experiment, illustrated in figure 1.3, indicate that it is effectively impossible to mask a change in the residue pitch produced by the high harmonics by a band of noise centred on the fundamental frequency.

A third and highly convincing demonstration that residue pitch is not the result of a quadratic difference tone at the fundamental frequency has been described by de Boer (1956). If the components of some harmonic signal which produces a residue pitch are each shifted arithmetically up or down in frequency, then the frequency difference between these components remains constant, so that this manipulation will not affect the frequency of a quadratic difference tone. De Boer observed that as the component frequencies were shifted, the residue pitch of such complexes also shifted, in the same direction as and in approximate proportion to the shift of the component frequencies; since this residue pitch changed despite the invariance of the frequency of any possible quadratic difference tone, it may be concluded that this pitch was not due to such a distortion product. (This and related experiments will be discussed in more detail below, pp27-30).

Ohm's principle that a component of frequency f Hz must necessarily be present if the corresponding pitch is to be heard is thus untenable. Schouten (1940) proposed an alternative hypothesis, that residue pitch was due, not to the presence of a simple component of the appropriate frequency, but to the

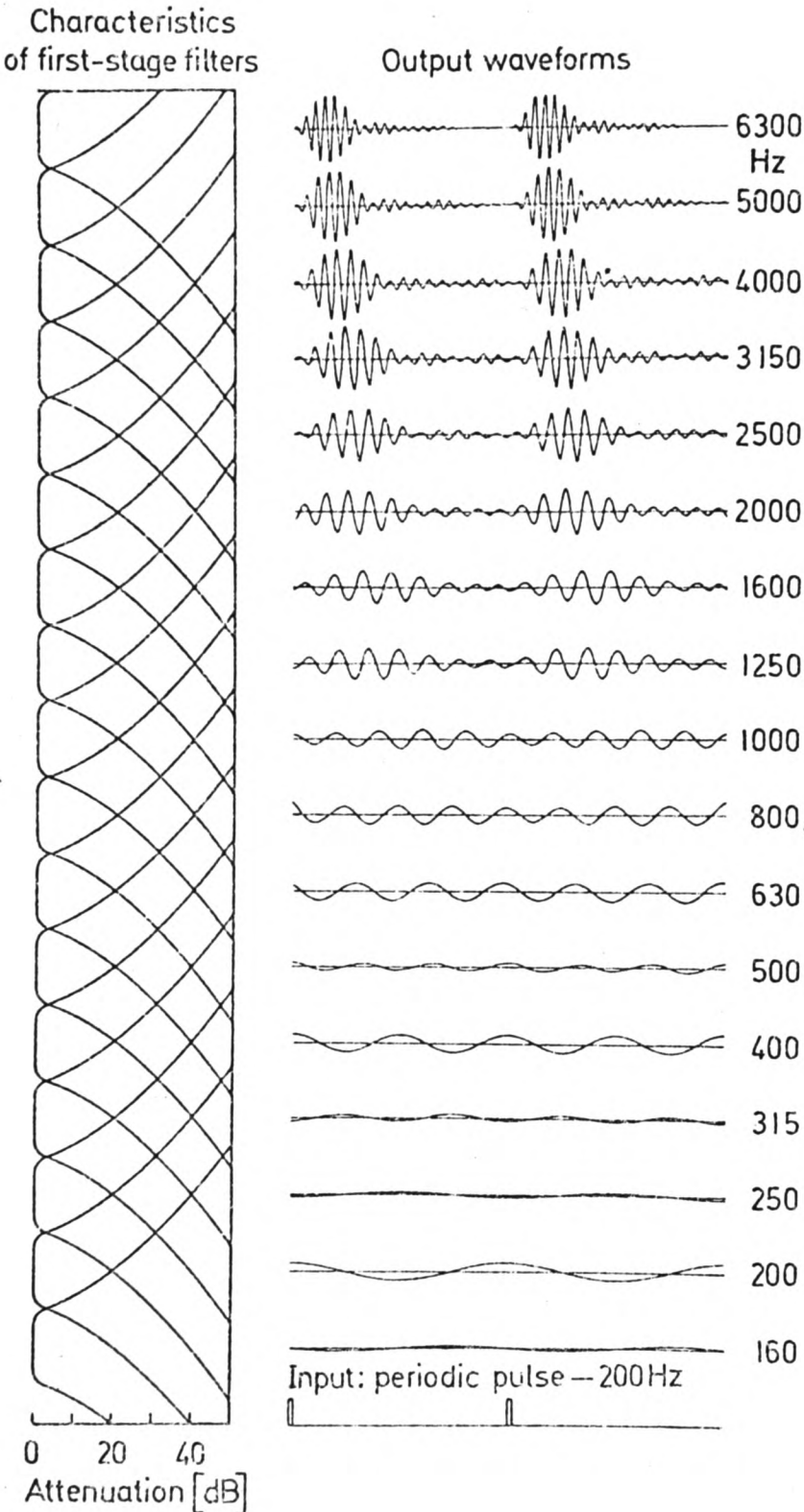


effectiveness of the periodicity of the waveform of the signal in determining pitch. Schouten's view did not deny the occurrence of auditory frequency analysis; rather he supposed that this analysis was of limited resolution, so that the higher harmonics of a complex signal were incompletely analysed and that the time-varying properties of the unanalysed or residual portion of the signal preserved the periodicity of the signal waveform.

Frequency Analysis and Residual Periodicity.

The assumption that auditory frequency analysis is of limited resolution is supported by a large body of evidence from both physiological studies (e.g. Békésy, 1960; Kiang, 1965; Rose, Hind, Anderson, and Brugge, 1971) and psychophysical investigations (e.g. Houtgast, 1974; Plomp, 1964; Plomp and Mimpen, 1968; Scharf, 1970), and it provides also a useful qualitative model of the behaviour of the first stages of the auditory system. Such a model, which approximates the frequency-selective behaviour of the basilar membrane (Békésy, 1960) consists of a series of frequency-selective elements, which may be represented by band-pass filters whose centre frequencies are distributed over the audible range, and whose bandwidths are approximately proportional to their centre frequencies. This model (Plomp, 1966; Schouten, 1940) and its behaviour when stimulated by a pulse-train, are illustrated in figure 1.4. The outputs of those filters centred on the lower harmonics of the complex signal are effectively sinusoids at the component frequencies, while the filters tuned to the higher

Figure 1.4. The output waveforms of a series of band-pass filters for a periodic pulse-train input of fundamental frequency 200 Hz. (After Plomp, 1966).

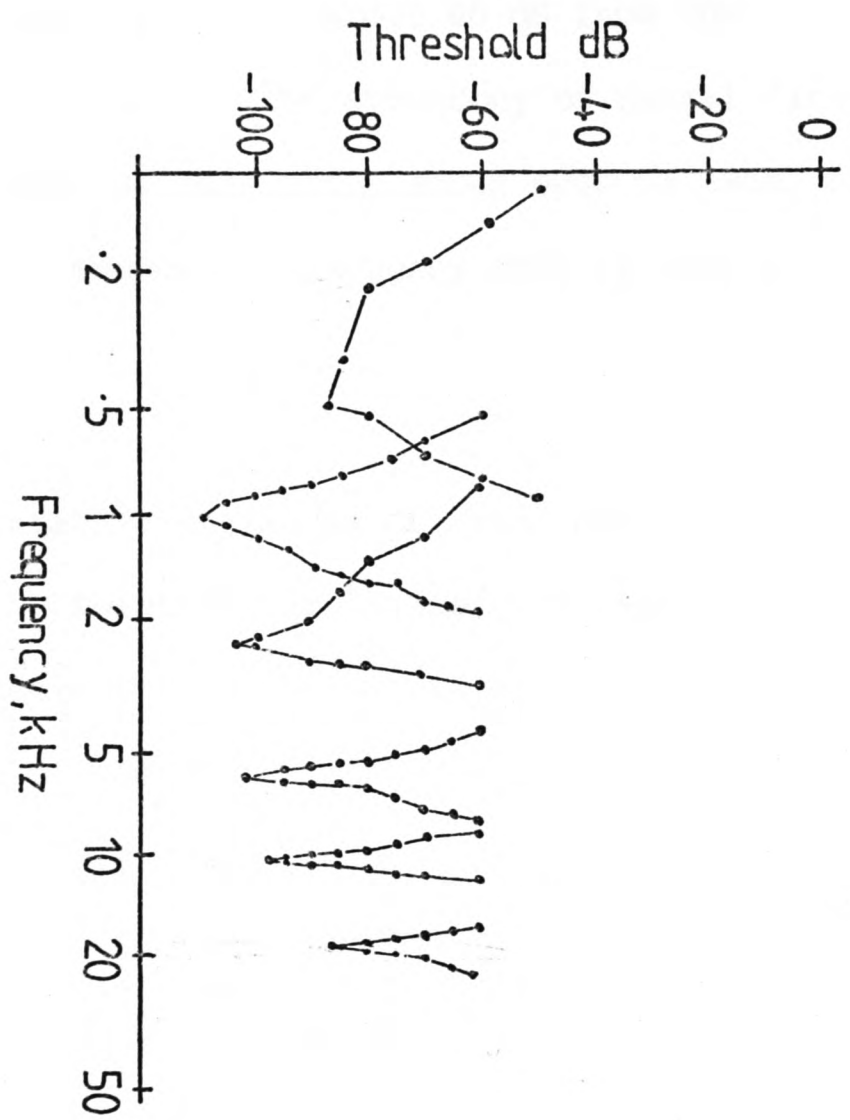


harmonics fail to resolve the simple components, and show a residual output which is similar to the sum of several adjacent components of the signal, and which preserves in its temporal structure the periodicity of the signal. Since the lower components are represented by sinusoidal outputs, their frequencies may, in principle, be measured either by reference to the period of the filter outputs, or by reference to the centre frequency of that filter in each local region which produces an output of the greatest amplitude, this local peak being equivalent to a displacement maximum at a particular place on the basilar membrane. In the case of the higher harmonics, the component frequencies cannot be clearly represented by either of these methods. The complex signal's periodicity is however well-preserved at the outputs of the filters tuned to the higher components, while little trace of this periodicity appears in the outputs of the filters responding to the lower components.

Auditory filters can be considered in two rather different ways. One view of the auditory filter is that a filter serves simply as part of a device for detecting energy at frequencies within the pass-band of the filter. Such a device would provide an output related to the energy within the pass-band, but would not discriminate between different signals which, after passing through the filter, have equivalent energies. If a signal affected two or more detectors, interactions between the detectors might, as Békésy (1960) suggested, allow a greater degree of discrimination. Such a system can only yield an estimate of the amplitude or power spectrum of the signal, and

all of the information contained in the signal's phase spectrum will be lost. An alternative view of the auditory filter, whose consequences are sometimes neglected, is that the output of a filter is some function of time which is related to both the energy of the signal within the pass-band of the filter and to the temporal properties of those components of the signal which fall within the pass-band. A series of devices of this sort permit the extraction of an estimate of the power spectrum of the signal, and, since the outputs of the filters are temporal representations of the signal, the phase spectrum of well resolved components will also be preserved. In the case of poorly resolved components, the temporal properties of the outputs of the filters will permit discrimination between different signals whose bandwidths are less than that of a single filter. This latter view of auditory frequency analysis is consistent with the firing-patterns of the auditory nerve, which are known to show both spatial and temporal coding of frequency. One source of evidence for spatial coding is a study of the auditory nerves of the cat by Kiang (1965). Kiang demonstrated that the threshold of an auditory nerve fibre is lowest at a particular stimulating frequency, and that the rate at which the fibre's threshold increases with the ratio between this 'best' frequency and the stimulus frequency is approximately constant for fibres with 'best' frequencies above 2000 Hz, while this rate decreases monotonically with frequency for fibres with 'best' frequencies below this. The results of this study are shown in figure 1.5; the limit of 2000 Hz may be species dependant, but the superior selectivity at higher frequencies is probably typical of all mammals. Temporal coding

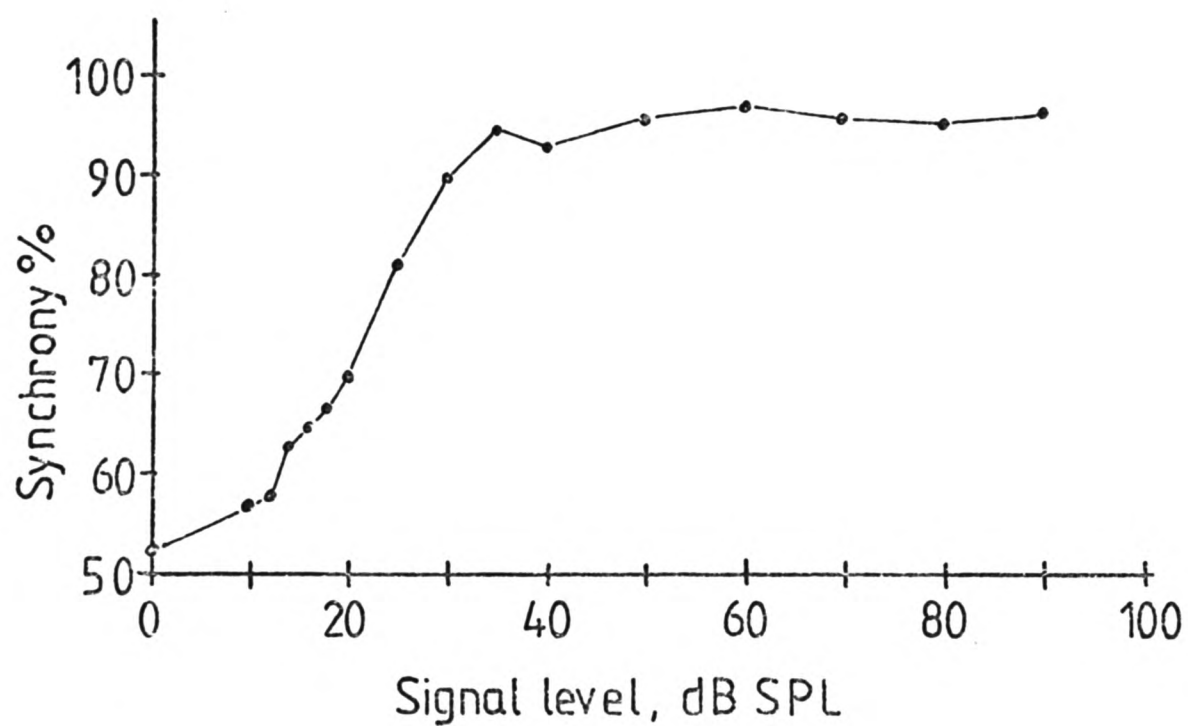
Figure 1.5 Neural tuning curves. The threshold of auditory nerve fibres as a function of frequency. (After Kiang, 1965).



of frequency is demonstrated by the synchrony of the firing pattern of the auditory nerve to peaks of the instantaneous amplitude of the stimulus. This phenomenon has been extensively studied by Rose and his associates (e.g. Rose, Brugge, Anderson, and Hind, 1967). The degree of synchrony depends in part upon the intensity of the stimulus, and continues to increase over a range of about 60 dB from the fibre's threshold (see figure 1.6). The synchrony of neural firing breaks down for sinusoidal stimuli above about 4000 Hz (Rose et al, 1967), so that a temporal frequency code is most effective at lower frequencies.

Quantitative estimates of the extent to which auditory frequency resolution is limited will be discussed below; here it is sufficient to note that if the input signal is composed of a series of harmonic components, then the lower components, which are widely spaced relative to the bandwidth of the filters which respond to these components, will be well resolved, while the higher components, spaced closely together relative to the bandwidths of the filters of higher centre frequencies, will not. Provided that other limits are not imposed in the subsequent transmission of this information, the periodicity of the signal should be better preserved for a signal containing high harmonics, since here the periodicity is carried by a single channel, and a theory of residue pitch perception which assumes that pitch is determined by the signal periodicity extracted as a result of such incomplete resolution must predict that residue pitch will be better perceived where some components are poorly resolved. A theory which supposes that

Figure 1.6 Neural synchrony as a function of the level of a sinusoidal signal. The signal frequency was equal to the 'best' frequency of the fibre, which was 4000 Hz. The measure of synchrony is the proportion of neural spikes occurring during the half-cycle of the stimulus to which the fibre is most responsive. (Rose et al, 1967).



residue pitch is based upon the use of spectral information, or upon the temporal coding of individual components, must, on the other hand, lead to the prediction that residue pitch will be better perceived where the components are well resolved.

The Perception of Residue Pitch as Spectral-Pattern Recognition.

A alternative view of the process leading to the perception of residue pitch is that the spectral pattern provided by the resolved components of a signal could be compared with an generalized internal representation of the spectrum of a periodic signal, which may be likened to a spectral template. If, for the sake of a simple graphic illustration, the spectrum of the signal, and the spectrum represented by such a template are each scaled logarithmically, then the template may be shifted along a frequency dimension and the two spectra directly compared by computing the correlation between the two. Such a comparison is shown in figure 1.7. If the fundamental component of the template is placed at the fundamental frequency of the signal, then the components of the signal's spectrum will each correspond to the frequency of one component of the spectral template, and the correlation will be maximal. If the fundamental frequency of the template is incorrectly chosen, then in some cases, the correlation between the template and the signal spectrum may still be fairly high, and a certain degree of ambiguity of pitch would be expected. Such ambiguity is indeed observed (Schouten, Ritsma, and Cardozo, 1962). The match shown in panel c) of figure 1.7 illustrates a case where

the correct fundamental frequency of 200 Hz has been selected, and the 5th, 6th and 7th harmonics of the spectral template match exactly the frequencies of the three components of the signal. The other two panels illustrate cases where the template has not been placed with its fundamental component at the same frequency as the fundamental of the signal. Although an approximate match is possible between the three components of the signal and - in panel a) - the 6th, 7th and 8th components of the template, or - in panel b) - the 4th, 5th and 6th components of the template, neither of these two matches are as well-fitting as that for the correctly chosen fundamental frequency. Well-fitting mis-matches of this sort are possible only if either the estimated spectrum of the signal or the spectral template, or both, are noisily represented; if both were represented by a line spectrum, then there could be no correlation between components whose frequencies did not match exactly. Such a matching process may be termed 'spectral-pattern recognition', and the underlying notion of a search for some well fitting estimate of fundamental frequency was suggested by de Boer (1956). Several models, formalised in a variety of ways, but generally analogous to the notion of a search for a well-fitting estimate of fundamental frequency, have been described, and will be discussed below.

A spectral-pattern recognition process of this sort does not make use of all of the spectral information which auditory frequency analysis is likely to provide. The model of auditory frequency analysis described above and illustrated in figure 1.4 permits the extraction from each well-resolved component of not

only its frequency, but also its amplitude and its phase relation with some other temporal referent. The above model of spectral-pattern recognition may be sensitive to the amplitude of a component to the extent that the precision of the estimation of the frequency of that component may deteriorate with decreasing amplitude, but the matching process is completely insensitive to the phase angles of the components, and indeed, sensitivity to the phase angles of the components of the signal would not provide a process intended for fundamental frequency estimation with any additional useful information.

Residue Pitch as Temporal Pattern Recognition.

A further alternative view of the processes underlying the perception of residue pitch might be that the pitch is estimated in the time domain, the information employed by the process being not, as Schouten has proposed, the periodicity of an unanalysed residue, but individual temporal representations of each well-resolved component. Auditory frequency analysis will allow such a representation, and it is clear from studies of the localization and lateralization of sound (Stevens and Newman, 1936) that, for component frequencies below about 2000 Hz, human observers are able to make use of the phase information which would be provided by such representations of components. Some process which was able to combine, either by addition or by some more complex operation, the temporal representations of each individual component would generate a representation of the whole signal whose temporal properties showed the same periodicity as that of the original signal; a process of this

sort will be termed a 'periodicity reconstruction' process. While a system of this sort may appear to be essentially phase sensitive, such sensitivity can only be expected where the necessary precision is attained in the temporal representation of each component.

This method of periodicity extraction differs from that proposed by Schouten in that the information employed by the process will, as in a spectral-pattern recognition mechanism, be transmitted through a number of independent channels, and each channel may contribute its own independent internal noise, whereas Schouten's theory permits the transmission of sufficient information through a single channel with only a single noise source. A second difference is that if periodicity is reconstructed from temporal information carried by a number of independent channels, then phase sensitivity should be stronger where individual components are well-resolved, since it will be these components whose individual temporal structure is best represented; the sensitivity of Schouten's theory to the relative phases of the components of a complex will however be greatest where the components are poorly resolved, since here it is more likely that several components will contribute to the temporal structure of the unanalysed residue. It is of course possible that a periodicity extraction mechanism based upon the recombination of temporal representations of the components of the signal might discard any relative phase information, and thus become phase-insensitive; if this were the case, such a mechanism would be generally equivalent to a mechanism employing the spectral cues embodied in the amplitude spectrum of the

signal. One plausible process by which relative phase information could be discarded is autocorrelation (Wiener 1948). Licklider (1951, 1959) has described a theory of the perception of residue pitch based upon a neural autocorrelator, and if this theory, which is discussed immediately below, is extended to allow temporally coded resolved components to contribute to a neural autocorrelation, it then embodies a phase-insensitive periodicity reconstruction process.

Licklider's Neural Autocorrelation Theory.

A development of Schouten's theoretical position was described by Licklider (1951, 1959). Schouten's view implies a many-to-many representation of pitch in that a variety of stimuli producing a single residue pitch may be represented in many places along the basilar membrane and in the associated array of nerve fibres, and stimuli having differing pitches may be represented in the same place. Licklider proposed a neural mechanism by which such a many-to-many representation could be transformed to a one-to-one representation, each pitch having its own neural receptor. The running autocorrelation $\overline{R}(\tau, t)$ over a time window t_1 to t_2 of a signal $f(t)$ may be calculated as;

$$\overline{R}(\tau, t) = \sum_{t=t_1}^{t_2} \sum_{\tau=t_1}^{t_2} f(t) \cdot f(t-\tau),$$

where τ is a delay interval. If $f(t)$ is a periodic signal embedded in noise, its autocorrelation is periodic over the period of the signal. As the duration of the autocorrelated signal is increased with respect to its period, the periodicity of the autocorrelation, unlike the periodicity of the signal

itself, is increasingly unaffected by the noise, permitting a more accurate estimate of the period. The autocorrelation of a signal also discards phase information, and resembles the waveform of a signal with the same power spectrum as the original signal, but with all the components in cosine phase. Licklider suggested that the residual periodicity of incompletely analysed complex sounds could be processed by a neural network which approximated an autocorrelator, and which ultimately generated a unique spatial representation of the residue pitch of such a sound. A schematic representation of such a neural network is shown in figure 1.8. Although Licklider's model makes use of the same stimulus information as Schouten's periodicity theory, the behavioural predictions of Licklider's model are rather different from those of Schouten's theory. One major difference between the two is that while the temporal structures of the signal waveform and of the unanalysed residue from an auditory frequency analysis of limited resolution are dependent upon the relative phases of the components of the signal, the autocorrelation function of the residue will be independent of relative phase. It is also quite conceivable that the input to such an autocorrelator would include temporal representations of a number of well-resolved sinusoidal components, so that while Schouten's theory proposes that only unanalysed components contribute to residue pitch, Licklider's model also allows for a contribution from completely analysed components. Probably because of the lack of any evidence at the time that lower harmonics played any part in the perception of residue pitch, Licklider's discussions of the model disregard this possibility, and implied that resolved

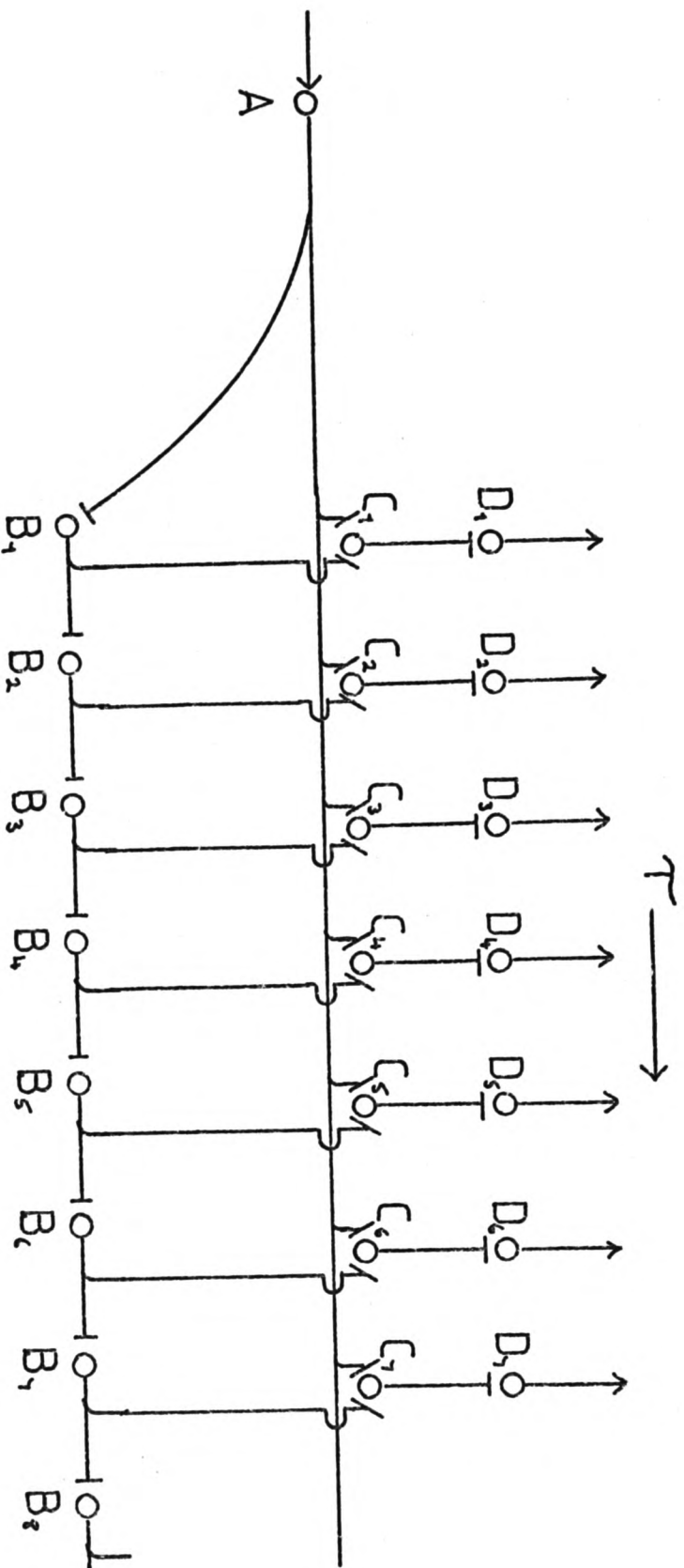


Figure 1.8. Licklider's neural autocorrelation model. The neural input, which is synchronized to the stimulus, is relayed from A to a series of delay lines, the neurons B. At each of the neurons C_i , the delayed signal is effective only when a spike coincides with a spike from the undelayed neural signal. If the delay of a given B neuron coincides with the periodicity of the signal, the associated C and D neurons will show a consistent response, so that the periodicity of the input is represented as a spatial distribution of activity along the series of neurons D_i . (From Licklider, 1951).

components were simply heard as individual pitches. The autocorrelation method of periodicity extraction has received more recent attention, and has been adopted by Wightman (1973b) as the basis for a model of pitch perception based upon spectral cues. This model is treated in detail below. It differs from Licklider's model in a major respect: The autocorrelation is not calculated purely in the time domain, but from an estimated power spectrum resulting from an auditory frequency analysis. Thus, Wightman's autocorrelation model only makes use of information from resolved spectral components and not from any residual periodicity. The neural network proposed by Licklider has been treated with scepticism by some neurophysiologists (e.g. Whitfield, 1970) because the neural delay lines required by the model have not been observed.

If the frequency coding present in the temporal firing pattern of the auditory nerve fibres is utilized for pitch perception, the extension of Licklider's theory to the use of information from resolved components has a considerable degree of explanatory power and is also functionally efficient given that a certain amount of noise is present in the neural channels. Such noise may either be external to the auditory system, as when acoustic noise is added to the signal, or it may be internal noise, which will inevitably arise from both the spontaneous activity of the nerve fibres and from the stochastic nature of neural time following (Rose et al 1967).

Consider a simple signal of two components which are sufficiently different in frequency for each to be

well-resolved: The outputs of two filters centered on the two components will provide the least noisy temporal representation of the signal. The best representation of periodicity from a single filter will be obtained from a filter positioned between the two components, and because the relatively fine resolution of such a filter will cause considerable attenuation of the two components, its output will have an unfavourable signal-to-noise ratio, and thus give a rather imprecise representation of the periodicity. Figure 1.9a) illustrates such a configuration of filters and components. In the case of a signal composed of two components closely spaced with respect to filter bandwidths, the converse will be true: In order to obtain a representation of each component alone, the two filters must be placed such that just one component falls within the pass-band of each, so that the filters are centred outside the two components. In this case, the desired components will be greatly attenuated, and thus very noisily represented. A single filter centered between the two components will, however, attenuate the two components only slightly, so that the periodicity of the sum of the two components here will be represented with a favourable signal-to-noise ratio. An illustration of this situation appears as figure 1.9b). This argument suggests a duplex theory of residue pitch (Licklider, 1951), in which the stimulus information responsible for residue pitch is that which is most precisely represented in the temporal pattern of the auditory nerve activity, and the extended version of Licklider's model described above could do just this. A test of such a theory against a theory proposing that only information derived from resolved frequency components is employed as a pitch cue thus

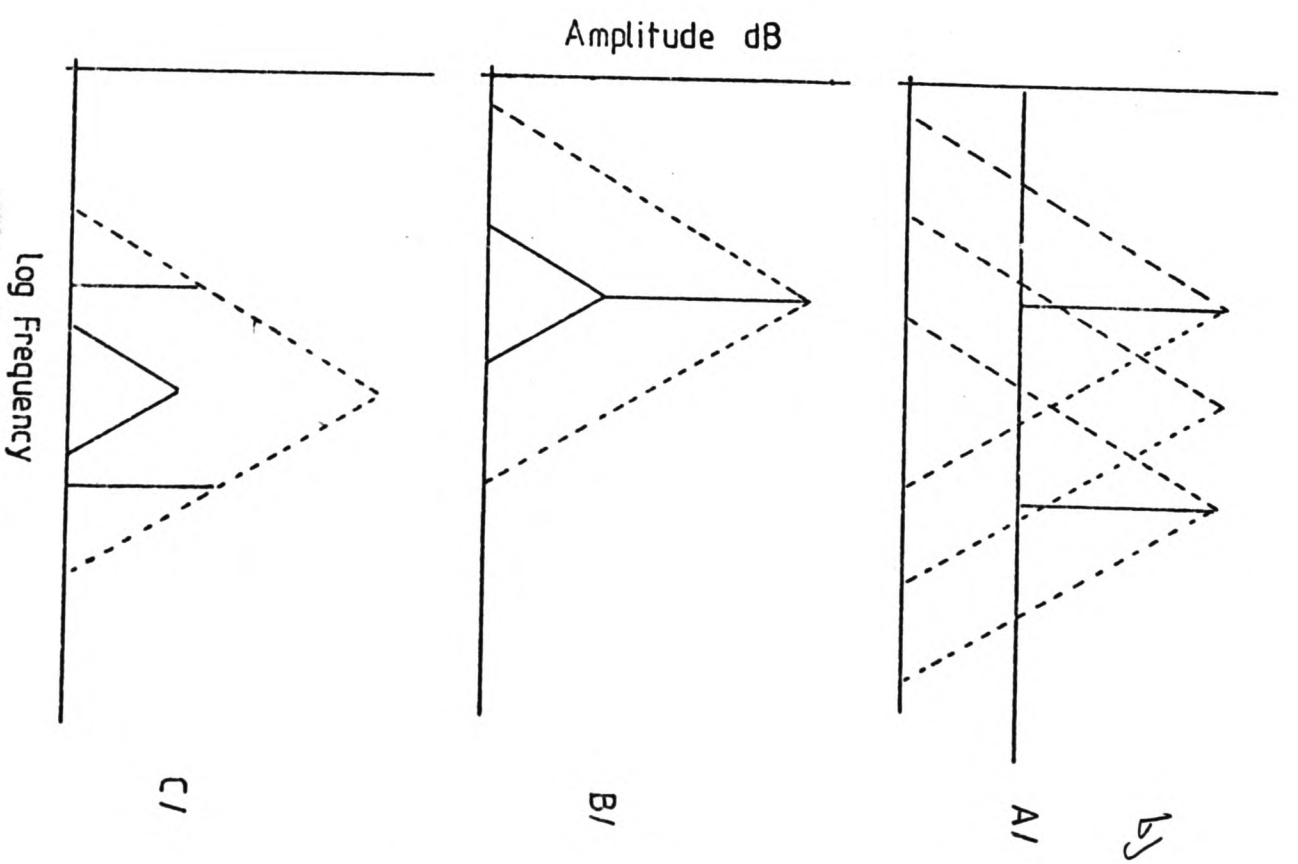
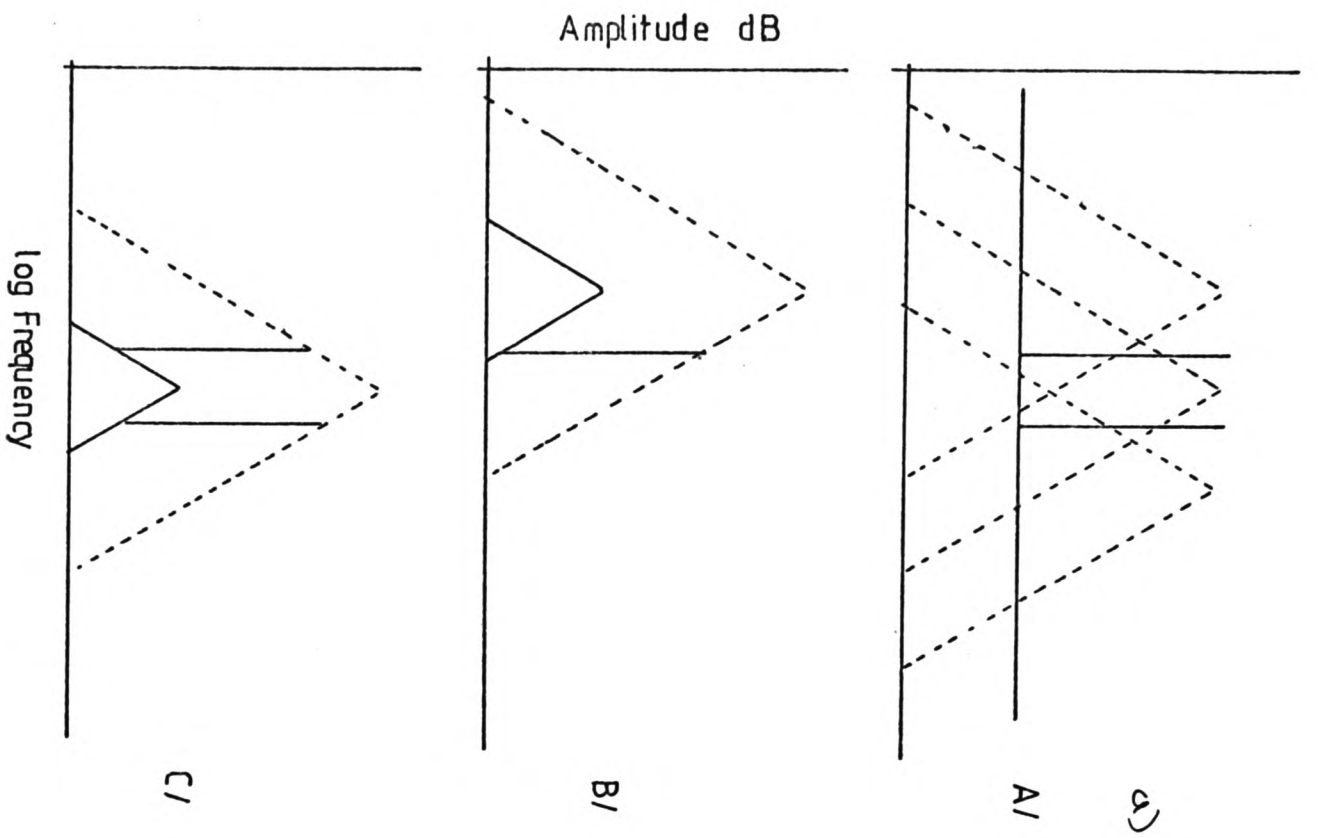


Figure 1.9 Panel a) represents the case of a signal whose two components are closely spaced relative to the resolution of the auditory filters (shown as triangular attenuation functions). Panel b) represents a signal whose components are well resolved. In each panel, A/ shows the signal and noise spectrum prior to filtering, while B/ and C/ show the signal and noise spectrum after filtering to optimize the representation of the lower component (B) and the periodicity (C).

depends greatly upon the availability of good estimates of the limits of frequency resolution as a function of frequency. This theoretical possibility is, however, the least parsimonious of those described here, and unless it can be shown that no single process theory can provide an adequate account of the data, a single process theory is to be preferred.

The Pitch Shift, Envelope or Fine-Structure?

Schouten's suggestion that the periodicity of the unanalysed residue is responsible for residue pitch is open to two possible interpretations. The important temporal property of the residual waveform may be either the true period, as measured by reference to the time interval between such corresponding points as successive major peaks in the temporal fine-structure, or it may be the envelope rate, a rather coarser estimate of periodicity, and one which, as will become clear shortly, does not always correspond to the true period.

De Boer's (1956) demonstration of the pitch shift of residue pitch, already mentioned in the context of the non-linear distortion hypothesis (p13), also permits a test of the appropriateness of the two representations of signal periodicity which might, if Schouten's view is correct, be responsible for residue pitch. The signals employed by de Boer were produced by the modulation of a sinusoidal carrier of frequency $f(\text{carr})$ by a complex signal containing components at the frequencies g , $2g$ and $3g$ Hz. After the addition of a component at the carrier frequency $f(\text{carr})$, the complex

contained seven components, $\underline{f(\text{carr})-3g}$, $\underline{f(\text{carr})-2g}$, $\underline{f(\text{carr})-g}$, $\underline{f(\text{carr})}$, $\underline{f(\text{carr})+g}$, $\underline{f(\text{carr})+2g}$, and $\underline{f(\text{carr})+3g}$ Hz. If $\underline{f(\text{carr})}$ is an integer multiple of \underline{g} , then the resulting signal is periodic, with a fundamental frequency of \underline{g} Hz, otherwise the signal may be termed 'quasi-periodic', since major peaks in the waveform occur at time-intervals close to but not equal to $\underline{1/g}$ sec. Examples of the waveforms of a periodic and two quasi-periodic signals generated by this method are illustrated in figure 1.10; the period of the periodic signal is $1/200$ sec, and the envelope also repeats in this time-interval; both quasi-periodic signals also show an envelope which repeats in $1/200$ sec, but the true periods of these signals are $1/20$ and $1/100$ sec.

The residue pitches produced by signals of this type were matched by comparison with the pitch of a similarly constructed, but always periodic signal, whose fundamental frequency was adjusted by the observer until the two pitches were judged to be equal. The results of this matching procedure, which are illustrated in figure 1.11, indicate that the residue pitch of a quasi-periodic signal does not correspond to the envelope rate of 200 Hz, but to a pitch slightly above or below this value. The alternative periodicity hypothesis, that the period of the signal is estimated from the temporal intervals between peaks in the fine-structure of the waveform is however, fairly consistent with the obtained matches. This result is, as Goldstein (1973) and Wightman (1973b) have shown, also consistent with the predictions of their respective spectral-pattern recognition models. Since the autocorrelation of these signals will be

Figure 1.10. Examples of the signals used by de Boer (1956) in his Pitch Shift experiment. Signal a) is truly periodic, and has a period of $1/200$ sec, and an envelope which repeats over its period. Signals b) and c) are quasi-periodic, having true periods of $1/20$ and $1/100$ sec respectively, these signals both have an envelope which repeats over $1/200$ sec. (from de Boer, 1976).

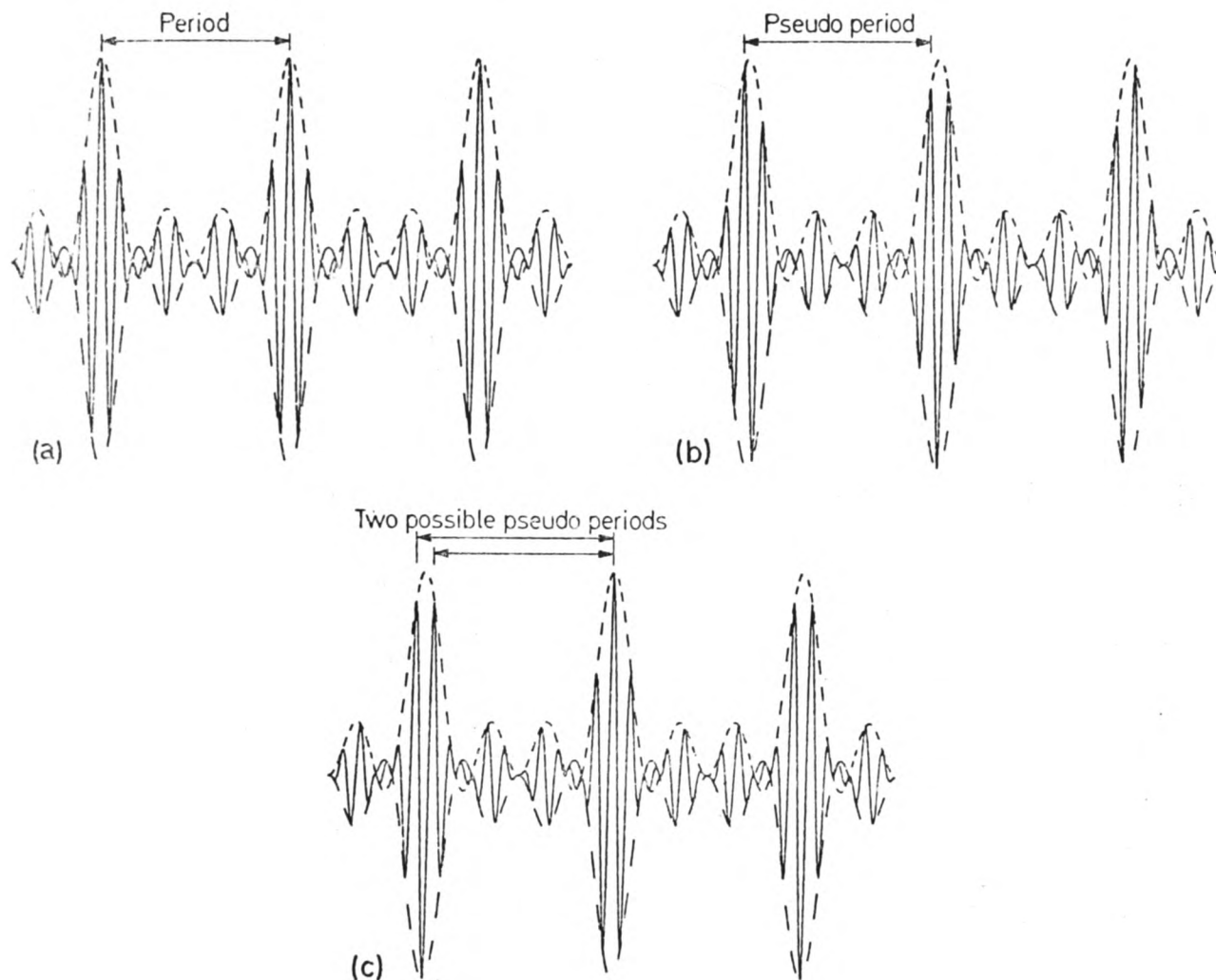
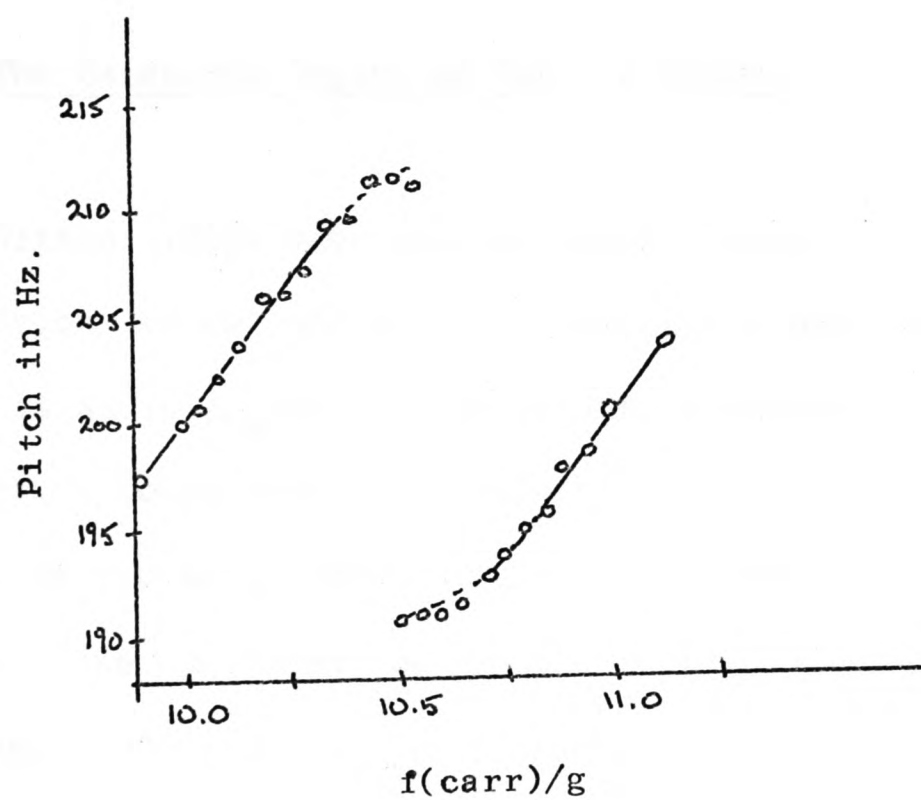


Figure 1.11 The pitch of inharmonic signals. The figure shows the fundamental frequency of a harmonic signal whose pitch matched that of the signal constructed by the amplitude-modulation of $f(\text{carr})$ by a sinusoid at g Hz. The pitch of the latter signal is ambiguous for $f(\text{carr})/g$ close to 10.5 (from de Boer, 1956).



similar to their waveform, Licklider's autocorrelation theory will also predict these results. Similar experiments investigating the pitch shift of the residue pitch formed by complexes of three components (Schouten et al, 1962), of two components (Smoorenburg, 1970), and of six and twelve components (Patterson, 1973; Patterson and Wightman, 1976) have also been reported. In all of these studies, the results were generally similar to those found by de Boer. This particular experimental manipulation also raises other interesting questions, and these studies will receive further attention below (pp 45-47).

The Existence Region of Residue Pitch.

Ritsma (1962) made an extensive study of the range of modulation rates and carrier frequencies under which a residue pitch is audible, and this so-called 'existence region' permits a useful comparison with evidence bearing upon the resolution limits of auditory frequency analysis. The test stimuli used by Ritsma were produced by sinusoidal amplitude modulation (SAM) (Littler, 1965) The carrier frequency $f(\text{carr})$ and the modulation frequency g were chosen so that $f(\text{carr})/g$ was always an integer, which will be designated as n , and which corresponds to the harmonic number of the carrier frequency in the complex signal. The modulation index m , which is equivalent to the ratio of the sum of the amplitudes of the two sidebands at frequencies $f(\text{carr})-g$ and $f(\text{carr})+g$ to the amplitude of the carrier frequency, was under the control of the observer, who was able to vary m between 0 and 1. The observer's task was to find, for a given pair of values of $f(\text{carr})$ and g , the minimum value of m

at which a residue pitch corresponding to \underline{g} Hz was just audible. The observer compared the pitch of the test signal with that of a reference signal which produced a clear residue pitch at \underline{g} Hz, the reference being a second SAM signal with $\underline{m}=1$, and the same values of \underline{g} and $\underline{f(carr)}$ as the test signal (in some cases, the pitch of such a reference signal was difficult to detect, and here $\underline{f(carr)}$ for the reference signal was reduced by $\underline{4g}$). Three observers made pitch judgements, \underline{g} being kept constant throughout each series of trials so that the observer was easily able to attend to pitches at \underline{g} Hz. The results of Ritsma's study are summarized in figure 1.12, which shows his estimates of the values of $\underline{f(carr)}$ and \underline{g} at which a given value of \underline{m} produces a noticeable residue pitch. The study did not include values of \underline{n} less than 4, because Ritsma considered that the resulting signals, which would contain a component at $\underline{2g}$ or \underline{g} , could be matched by reference to the pitch of this component rather than to the residue pitch of the complex. The results indicate that residue pitch is only audible when $\underline{f(carr)}$ is less than 5000 Hz, the exact limit being dependent upon \underline{g} , and that values of \underline{n} greater than 20 also fail to produce a clear residue pitch. For a given value of \underline{g} , the minimum value of \underline{m} can be seen to increase as $\underline{f(carr)}$ increases, so that as the relative spacing between the components decreases, and the components become more difficult to resolve, the necessary level of the sidebands is increased. These results were obtained for complexes presented at about 35 dB above the threshold of audibility; Ritsma also reported that presentation levels between 20 and 50 dB above threshold produced similar results.

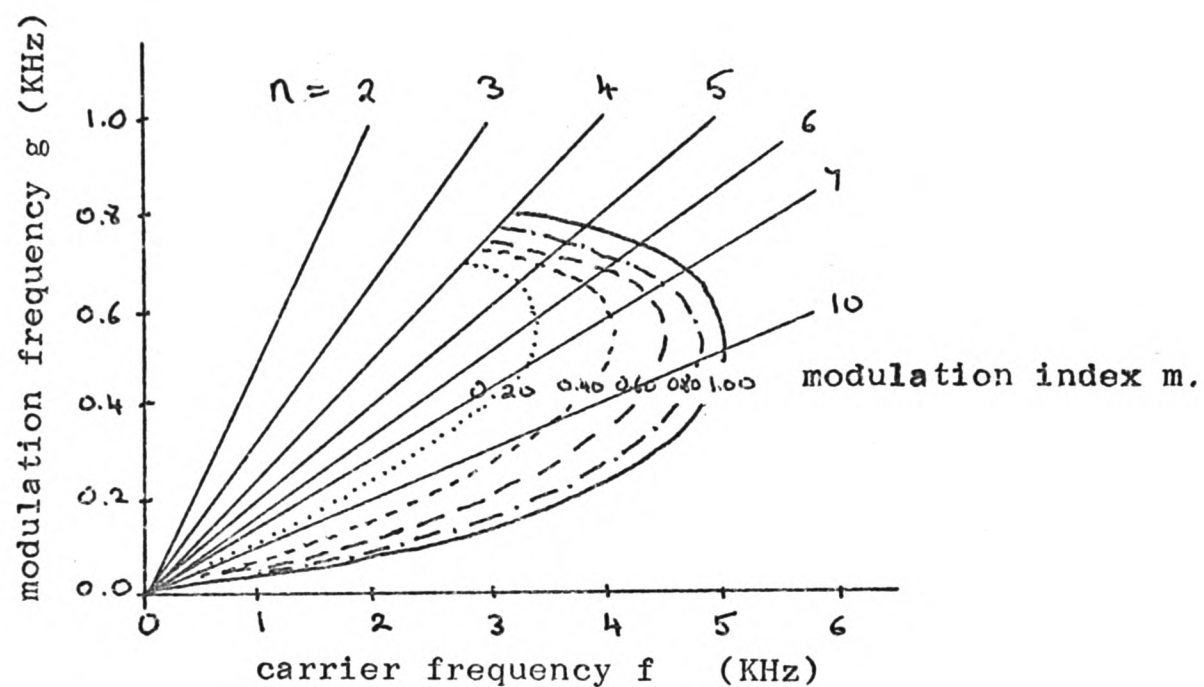
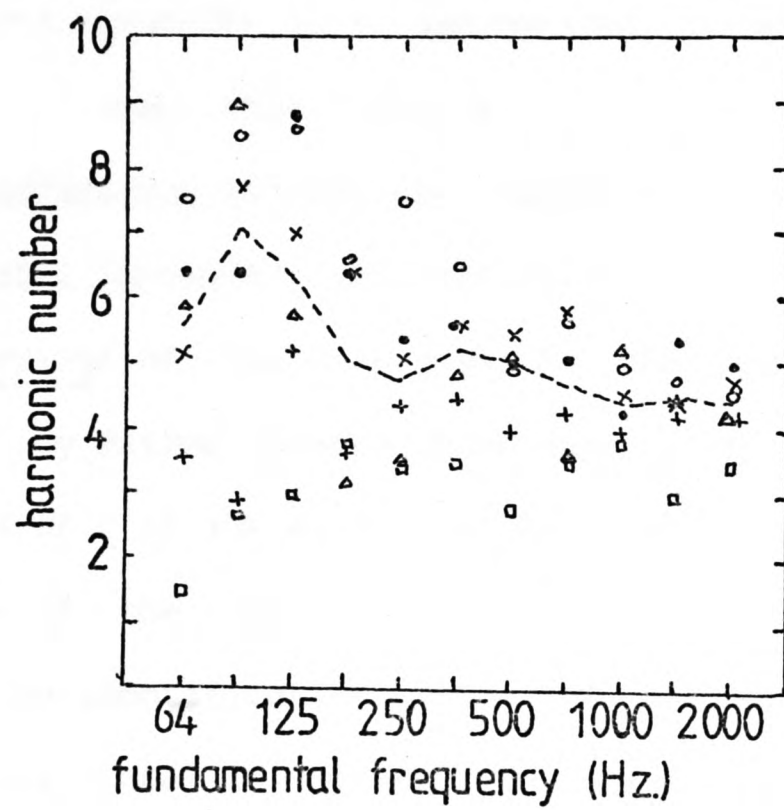


Figure 1.12 The existence region of the residue pitch of amplitude-modulated signals (Ritsma, 1962). The curves represent the boundaries of this region for various values of the modulation index m .

Frequency Resolution and the Existence Region.

The critical bandwidth (Scharf, 1970), has conventionally been accepted as an estimate of auditory frequency selectivity. According to Scharf's review, the estimated critical bandwidth is typically about 15% of the centre frequency for frequencies above 500 Hz. If the frequency separation between two components of a complex signal is smaller than the appropriate critical bandwidth, a critical band detector centred between the two components will respond to both, and the residual periodicity of their sum should be represented in the temporal discharge pattern of the auditory nerve. If the component separation is greater than the critical bandwidth, it should be possible to centre two critical band filters on the two component frequencies, and to estimate the frequency of each individually. Plomp (1964) and Plomp and Mimpen (1968) describe two studies which provide a direct estimate of the minimum frequency separation at which the components of harmonic twelve-component and two-component complex signals are identifiable. The results of these studies are shown in figure 1.13, where the harmonic number of the highest identifiable component of these complexes is shown as a function of the fundamental frequency of the signal; the frequency separation between the components is of course equal to the fundamental frequency. From these results, it appears that for fundamental frequencies between 100 and 300 Hz, no more than the first eight harmonics of a twelve-component complex, can be identified. In the case of a two-component complex, however, harmonics up to the 10th or 11th may be identified. The difference between these

Figure 1.13 Harmonic number of the highest identifiable component of a 12-component harmonic complex as a function of fundamental frequency. The different symbols represent the performance of the six individual observers, while the dotted line indicates the group mean. (Plomp and Mimpen, 1968).



signals seems likely to be due to the lesser degree of masking by adjacent components with the two-component signal.

If residue pitch were to be based upon the integrity of signal periodicity as preserved in the time-course of the output of frequency-selective elements whose bandwidth is greater than the spacing between adjacent components of the signal, and if the critical bandwidth is an appropriate measure of frequency selectivity, then this pitch should become harder to identify once the difference between the component frequencies exceeds the critical bandwidth, and this limit will arise for values of n ($n=f(\text{carr})/g$) of less than about 6. The existence region as delimited by Ritsma clearly shows that residue pitch is audible for values of n as low as 4, and this result is hard to explain in terms of the preservation of signal periodicity. If, however, the identification of the frequencies of the components of a complex signal is necessary for the perception of a residue pitch, as would be expected on the basis of spectral-pattern-recognition or periodicity reconstruction models, then residue pitch should become inaudible once the differences between the frequencies of adjacent components fall below about half of the critical bandwidth, that is, for values of n greater than about 10 to 12. Ritsma's results also belie this expectation, although the more general prediction of these latter models, that the perception of residue pitch should become more difficult as n increases, is supported by Ritsma's results. The absence of any residue pitch from complexes where $f(\text{carr})$ was 5000 Hz or greater is consistent with the breakdown of the relation between the firing rate of eighth-nerve fibres

and stimulus frequency at about that frequency (Rose et al 1967). This observation is a demonstration that temporal coding plays a part in the perception of residue pitch, since the estimated critical bandwidth ^{above 5000 Hz} increases roughly in proportion to frequency (Scharf, 1970), so that in this frequency region a spectral representation is likely to be no coarser than that available with lower component frequencies.

The Relative Importance of Low and High Harmonics; Spectral Dominance.

The contrasting predictions of periodicity and spectral-pattern recognition theories of residue pitch perception have already been mentioned. A theory based upon the use of the periodicity of an unanalysed residue leads to the prediction that the poorly resolved higher harmonics will determine residue pitch, while any theory based upon the use of information derived from the individual components will lead to the prediction that the better resolved lower components will dominate the perceived pitch. These two predictions may be directly contrasted by obtaining estimates of the pitch of complex signals containing both high and low components, and so constructed that the higher components presented alone would produce a different pitch from that produced by the lower components. Ritsma (1967) and Plomp (1967) each carried out an investigation of the dominance of the high and low components of such signals. The test signal used by Plomp was:-

$$f(t) = \sum_{n=1}^m a_n \cos[2\pi n(0.9f)t] + \sum_{n=m+1}^{12} a_n \cos[2\pi n(1.1f)t],$$

and the pitch of this signal was compared with that of a reference signal:-

$$f(t) = \sum_{n=1}^{12} a_n \cos[2\pi n \cdot f t].$$

The two signals were presented successively for 200 msec each, and the observer's task was to judge which of the two signals was higher in pitch. Plomp employed values of f from 125 to 2000 Hz, and values of m of 1, 2, 3 and 4 where m is the harmonic number of the highest frequency component with the fundamental frequency $0.9f$. The signals were presented at about 60 dB SPL, and the relative levels a_n of each component were either equal, or decreased in inverse proportion to n . The results of this study, averaged over 14 observers, appear as figure 1.14. For $m=4$, the majority of the pitch judgements were based upon the pitch produced by the lower components at all values of f , and it may be concluded from this result that harmonics above the 4th are of lesser importance in the determination of the pitch of a complex tone than the lower components. Since the fundamental component was present in this test signal, it might be supposed that it was this component which led to the dominance of the lower components, but the pitch judgements for $m=1$ suggest that, at least for $f < 1000$ Hz, the fundamental component is remarkably ineffective in determining pitch.

Ritsma's study of the range of dominance employed an essentially similar method. Here observers were asked to compare the pitch of two test signals A and B, constructed as follows:-

$$A. f(t) = \sum_{n=1}^m a_n \cos[2\pi n \cdot f t] + \sum_{n=m+1}^{\infty} a_n \cos[2\pi n(1.03f) t],$$

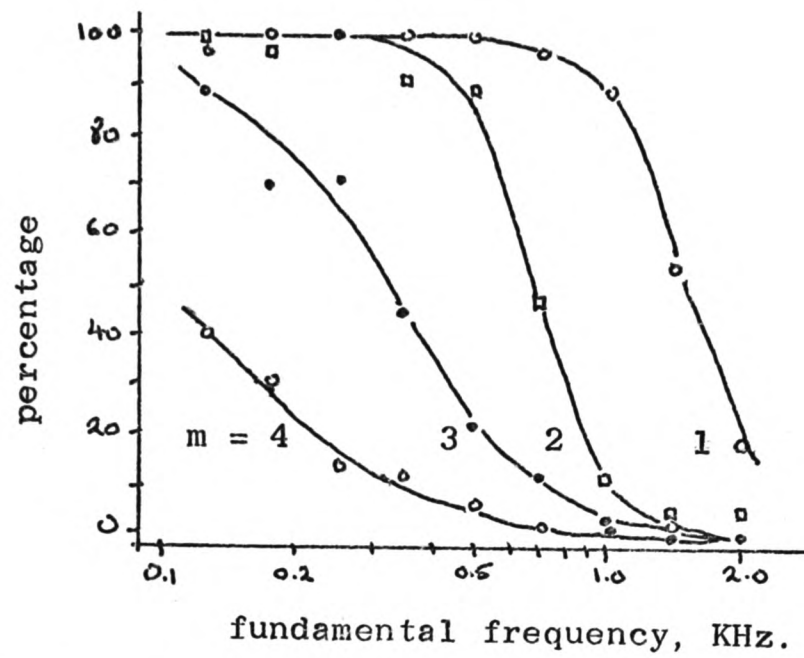


Figure 1.14 Percentage of trials on which the pitch of a complex sound was judged to rise when the $(m+1)$ th and higher harmonics were increased by 10%, and the m th and lower harmonics decreased by 10%. (Plomp, 1967).

$$B. f(t) = \sum_{n=1}^m a_n \cos[2\pi n(1.03f)t] + \sum_{n=m+1}^{\infty} a_n \cos[2\pi n.ft].$$

Signal A was produced by the addition of two pulse trains, one of fundamental frequency f , low-pass filtered at a cut-off frequency $\underline{f_c}$ ($\underline{f_c} = \underline{f} \cdot (2m+1)/2$), and one of fundamental frequency $1.03f$, high-pass filtered at the same cut-off frequency. Signal B was similarly constructed, the fundamental frequencies of the low-pass filtered and the high-pass filtered pulse trains being reversed. The two signals were presented successively for 200 msec each, with an inter-signal-interval of 300 msec; each signal was about 40 dB above audibility. The percentage of pitch judgements based upon the pitch expected from the higher components of the test signal for $\underline{f} = 200$ Hz is shown as a function of the cut-off frequency $\underline{f_c}$ for each of 4 observers in figure 1.15. These results are consistent with those obtained by Plomp. For $\underline{f_c} > 1200$ Hz ($\underline{m} > 6$), all four observers based the majority of their judgements on the information provided by the lower harmonics, while for $\underline{f_c} < 600$ ($\underline{m} < 3$), the higher harmonics were dominant. Since these results indicate a rather sharply defined dominance region lying between the 3rd and 6th harmonics, Ritsma explored the dominance of these components in more detail. This region of dominance coincides approximately with the frequency region over which the ear is most sensitive (Fletcher and Munson, 1933; Robinson and Dadson, 1956), so that it may be the case that this greater sensitivity is at least partially responsible for the dominance effects; Ritsma has however briefly reported other data which suggest that the lower harmonics continue to dominate in other frequency regions (Plomp and Smoorenburg, 1970). Plomp's study also bears upon this question, since he employed a range of

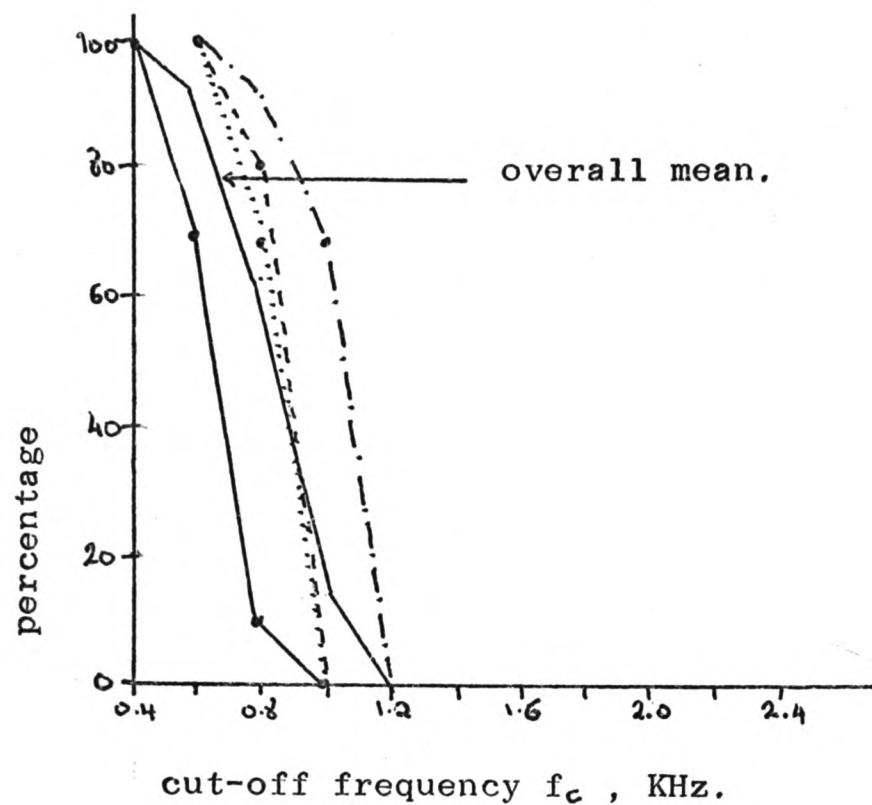


Figure 1.15 Percentage of trails on which pitch judgements were made in correspondence with the frequency change of the components above the cut-off frequency f_c . Results for 4 observers and mean. (Ritsma, 1967).

fundamental frequencies, and found that, for fundamental frequencies below about 1000 Hz, the frequency of the dominant component increased approximately proportionally with fundamental frequency, while for fundamental frequencies above 1000 Hz the first or second harmonic was dominant. A further demonstration that greater sensitivity is not responsible for the dominance region is provided by Ritsma's finding that the dominance of the lower components was independent of the relative presentation levels of the higher and the lower components as long as the lower harmonics were at least 10 dB above threshold: This result was observed for levels of the higher components up to 50 dB above threshold. Thus, Ritsma concluded that at least for complexes of fundamental frequencies between 100 and 400 Hz, the harmonics between the 3rd and the 6th are dominant in the determination of residue pitch.

Whatever mechanism may be responsible for the perception of residue pitch, it is likely that the better resolved lower components of a complex signal are more effective in determining pitch than the poorly resolved higher components. This conclusion, based not only upon the studies of dominance reported by Plomp and by Ritsma, but also upon Ritsma's (1962) study of the existence region, must favour a theory based upon the use of information derived from individual components rather than a theory based upon the preservation of the periodicity of the signal as a result of incomplete frequency analysis, although the possibility remains that the latter information is used where no other is available. Plomp's finding that the fundamental component of a complex sound is not dominant is

particularly interesting, because it suggests that the mechanisms responsible for the perception of residue pitch operate and dominate our perception of pitch even when the fundamental component is present.

The Effects of Relative Phase.

A further method of assessing the validity of these various theoretical viewpoints involves the manipulation of the relative phases of the components of a periodic complex. If such a signal has three or more components, then a change in the phase relations between the components can result in a pronounced change in the temporal fine-structure of the waveform, and would be expected to affect the operation of a periodicity extraction mechanism, this being likely both for Schouten's periodicity theory and for a theory which proposes that signal periodicity is reconstructed from a phase-sensitive temporal representation of individual components. Licklider's (1959) temporal autocorrelation theory discards phase information, so that this theory, despite being an extension of Schouten's periodicity theory, would not predict that relative phase will affect residue pitch. An amplitude or power spectrum is unaffected by changes of phase, so that a spectral-pattern-recognition system should be unaffected by this manipulation.

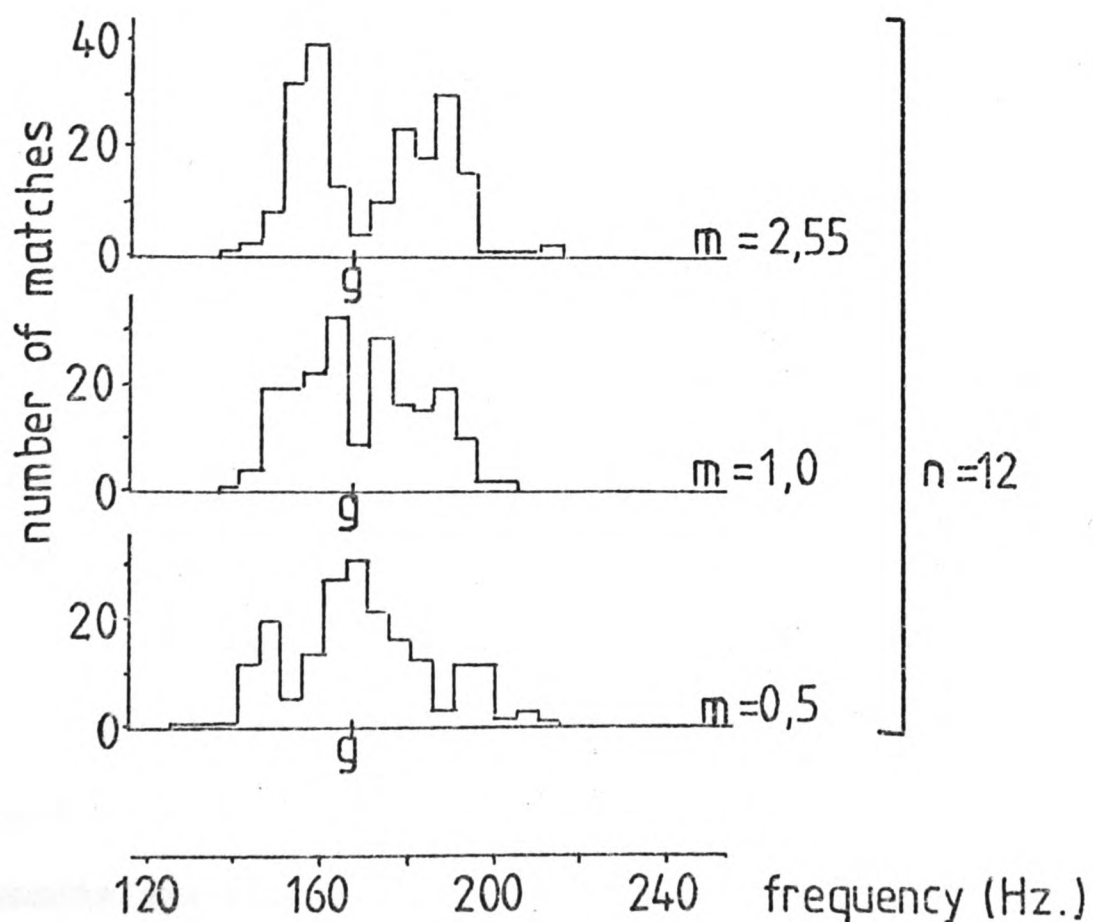
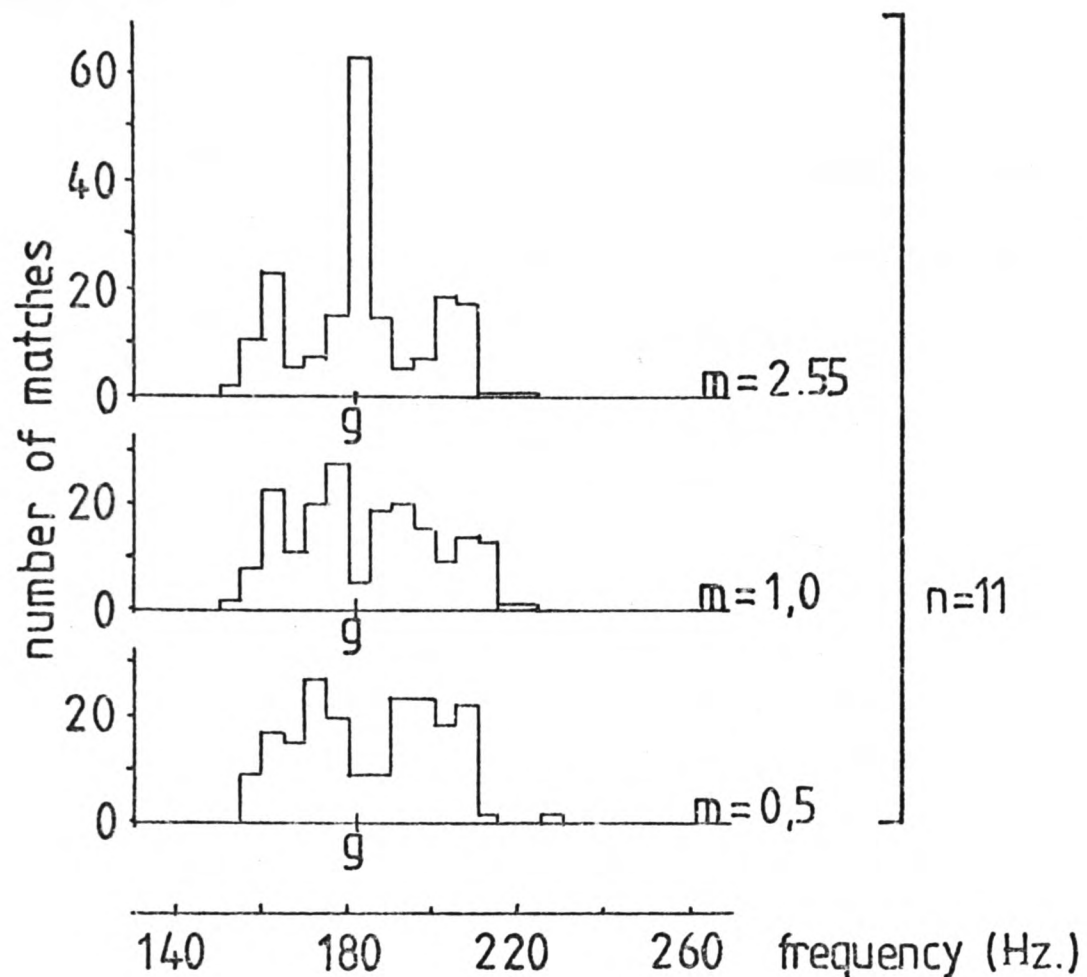
If the centre component of a three-component sinusoidally amplitude modulated (SAM) signal is shifted in phase by 90° , the envelope of the resulting waveform is rather flatter than that of the original SAM signal, and also tends to show peaks at

twice the modulation frequency \underline{g} . Such a signal, because of the resemblance of its waveform to that of a frequency modulated signal, is termed a 'quasi-frequency-modulated' (QFM) signal. The time interval between the major peaks in the SAM waveforms is the same as the period, or $1/\underline{g}$ sec; the waveforms of the QFM signals however, show secondary peaks whose instantaneous amplitudes are almost as great as those of the major peaks. A mechanism based upon the measurement of the signal periodicity, and subject to some random noise in the representation of the signal, would thus be likely, at least on occasion, to select time intervals other than $1/\underline{g}$ sec as estimates of the period of the QFM signal. The QFM signal with $\underline{n} = 12$ (where $\underline{n} = \underline{f}(\text{carr})/\underline{g}$) shows two secondary peaks, occurring approximately $0.45 \times 1/\underline{g}$ and $0.54 \times 1/\underline{g}$ sec before and after the major peaks, and corresponding to expected pitches of about $2.18\underline{g}$ and $1.84\underline{g}$ Hz. The QFM signal with $\underline{n}=11$ shows only one secondary peak, occurring approximately $0.5 \times 1/\underline{g}$ sec after the major peak, so that with this signal, a pitch of $2\underline{g}$ Hz may be expected. The SAM signals also show secondary peaks in their waveforms, but these all lie close to the major peaks, so that any error in the estimation of the period of the waveform is likely to result in a value close to $1/\underline{g}$ sec. The waveforms for $\underline{n}=11$ and $\underline{n}=12$ exemplify the features of such waveforms for odd and even values of \underline{n} (where $\underline{n} > 2$).

Ritsma and Engel (1964) investigated the perceived pitch of QFM signals with a constant carrier frequency of 2000 Hz, as a function of the modulating frequency \underline{g} , and \underline{m} , the modulation index, but they did not present any data for the comparable SAM

signals. Wightman (1973a) has reported pitch matching data for the same QFM complexes and for the comparable SAM signals. In both studies the same experimental method was employed: The observers were requested to adjust the fundamental frequency of a harmonic SAM signal to match the pitch of the test signal. All of the test signals were produced by the modulation of a constant carrier frequency of 2000 Hz by one of four modulating frequencies, the integer $\underline{n=f(carr)/g}$ being 10, 11, 12 or 13. The signals were presented at about 35 dB above threshold. Ritsma and Engel forced their observers to match the pitch of the test signal either in a region around \underline{g} Hz, or in a region around $\underline{2g}$ Hz. In Ritsma and Engel's study, the obtained distributions of matches in these two regions differed only in that the fundamental frequencies of the selected matching signals were an octave higher when in the region of $\underline{2g}$, and the results for $\underline{n=11}$ and $\underline{n=12}$ shown in figure 1.16 are the averages of these two sets of matches. The most striking feature of Ritsma and Engel's data is the ^{small number} of matches at \underline{g} or $\underline{2g}$ Hz for complexes where \underline{n} is 10 or 12; most matches to these QFM signals were at values either above or below \underline{g} or $\underline{2g}$ Hz, in accordance with the time intervals between major and secondary peaks in the waveform. Wightman's results, which are not illustrated, show no consistent differences between the distribution of pitch matches between the SAM and QFM signals, and matches at \underline{g} or $\underline{2g}$ were observed for all of the signals. The published details of these two studies suggest that apart from the use of different observers, the procedures were identical, so that the interpretation of the conflict between these results is problematical. Fortunately further data are

Figure 1.16 Distribution of pitch matches to the QFM signal. The upper panel illustrates the distribution of matches for a signal whose fundamental frequency \underline{g} was 181 Hz, at three values of the modulation index \underline{m} . The lower panel presents the distribution for a signal where $\underline{g} = 167$. Matches to pitches corresponding to frequencies greater than 250 Hz, are plotted at half of the matched frequency. (After Ritsma and Engel, 1964).

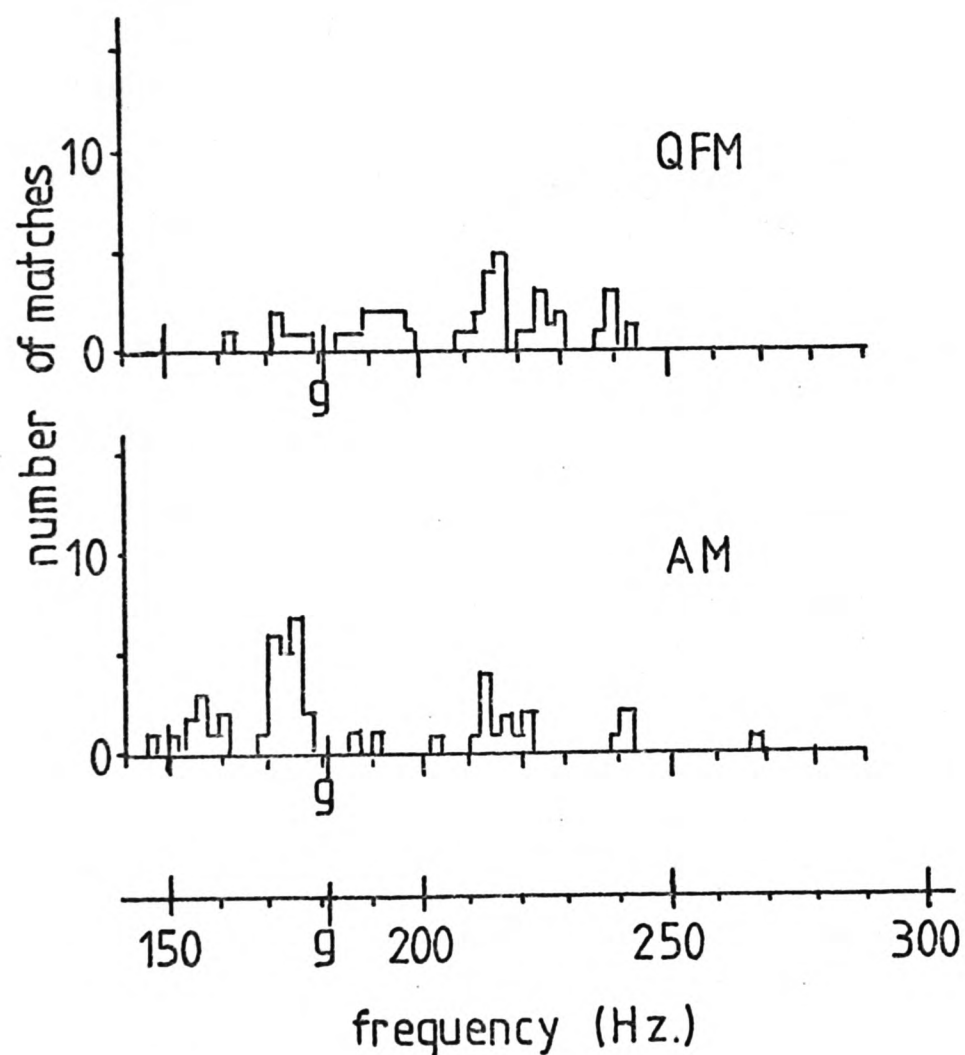


also available (Moore, 1977; Note 1), and this third study, which again employed essentially the same method, provides confirmation of the existence of the effects described by Ritsma and Engel, at least for some observers.

Moore was aware of the possible confounding role of aural distortion products (see below), and presented the above signals in a background of low-pass filtered noise so as to mask any possible distortion products at frequencies below those of the three signal components, but not the components themselves. The pitch matches of two of the six observers were apparently unaffected by the relative phase of the centre component, but the matches of the other four observers, at least in some of the experimental conditions, were so affected; figure 1.17 displays an example of the phase-dependent pitch matches produced by one of the observers.

Ritsma and Engel (1964) also investigated, using the same procedure, the pitch of QFM signals constructed with values of $n < 9$. These signals showed a prominent and unambiguous pitch of g Hz, irrespective of whether n was odd or even. This result may appear to be inconsistent with Schouten's notion that the temporal fine-structure of the signal provides the cue for residue pitch, but as Ritsma and Engel noted, the temporal fine-structure as seen at the output of auditory frequency analysis will be affected as a consequence of filtering. Where components are relatively widely spaced, as will be the case for lower values of n , then periodicity is likely to be extracted from the output of a filter responding to only two rather than

Figure 1.17 This figure gives one example of the relative phase dependence of pitch found by Moore (1977). The fundamental frequency g of both the AM and the QFM complexes was 181 Hz. The histogram shows the number of matches with a sinusoidal matching signal at frequencies between 145 and 300 Hz. (A few matches were also recorded at frequencies above 350 Hz., and these are not shown). The components of the complex signals were presented at 10 dB SL, and a low-pass filtered noise was also present. The noise was filtered at a cut-off frequency of $2000 - 2g$ Hz., with a slope of 60 dB/octave. The level of the noise was just sufficient to mask a 10 dB pure tone at $2000 - 2g$ Hz.



to three components, and the temporal fine-structure of the waveform of a two-component complex is very little affected by the relative phases of the components. If periodicity were to be derived not from the unanalysed residue, but by reconstruction, then the reconstructed temporal representation of the signal should, if phase information is not discarded, be dependent upon the relative phases of the components in the same manner as the waveform of the original signal. Because of the superior resolution of auditory frequency analysis for the more widely spaced lower harmonics, the temporal representation of these components will be more accurate than that for higher harmonics, and with such a theory therefore, phase sensitivity would be expected to be stronger for complexes composed of lower harmonics than for those composed of high harmonics, but it is clear that Ritsma and Engel's data fail to support this prediction.

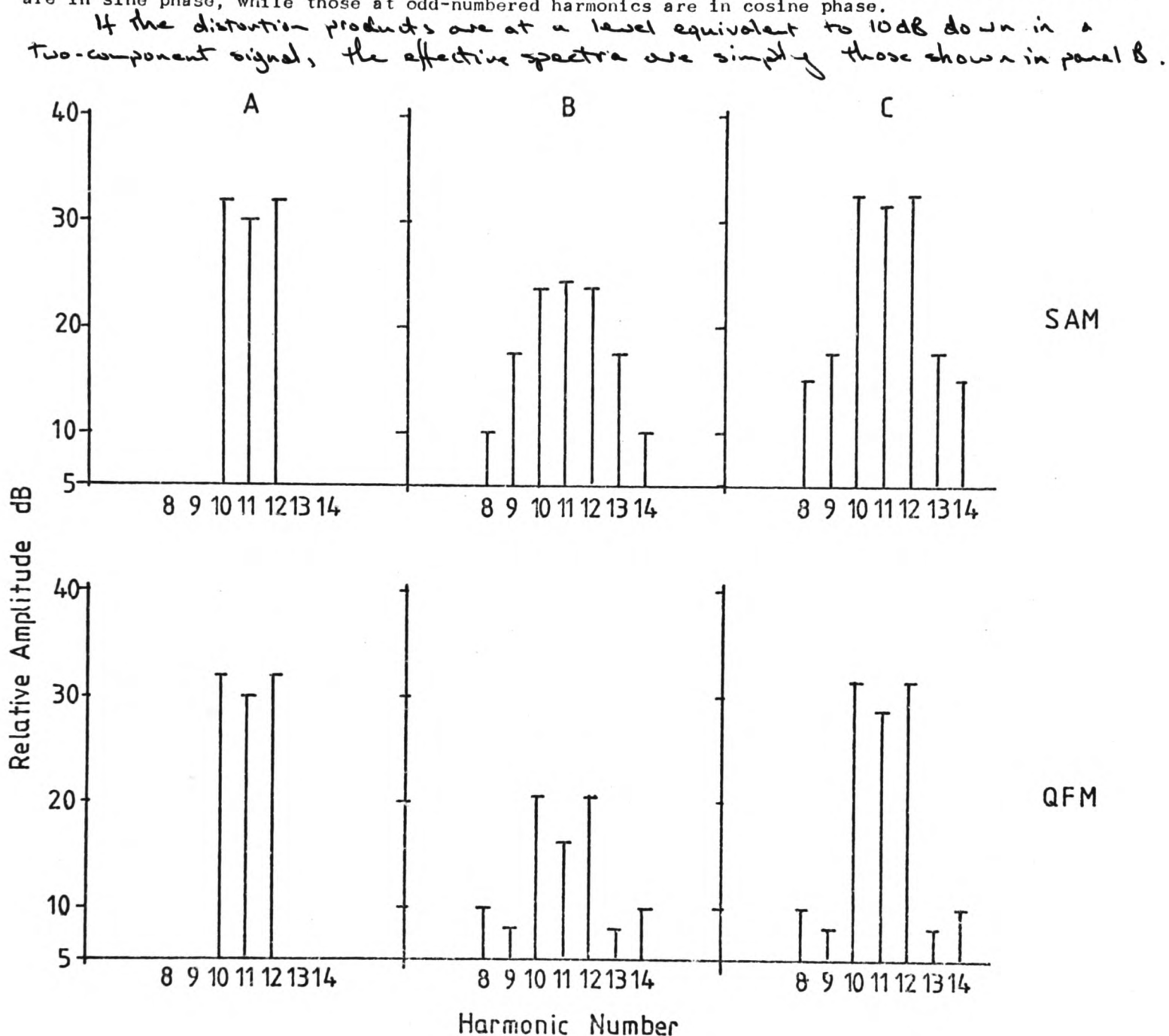
Relative Phase Effects and Auditory Nonlinearities.

Despite Wightman's failure to replicate Ritsma and Engel's data on the pitch of QFM complexes, these results, reinforced by Moore's (1977) replication, still pose a difficulty for theories of pitch perception based upon the use of information provided by resolved components. Since variation in the relative phases of the components of a complex signal has no effect upon its amplitude spectrum, such theories have no means of explaining Ritsma and Engel's results. One additional assumption, that significant nonlinearities occur in the peripheral auditory apparatus, may lead to a resolution of this conflict for

theories based upon the use of spectral information. The idea that the quadratic difference tone $f_2 - f_1$ was responsible for the perception of residue pitch has already been rejected, but the role of cubic and other higher order distortion products may be more important. The available evidence on the audibility of cubic distortion products (Goldstein, 1970; Plomp, 1965; Smoorenburg, 1972a, 1972b, 1974) suggests that these additional components are likely to be present to a significant degree with the signals employed by Ritsma and Engel and by Wightman, so that it is plausible that such a nonlinearity may have had a significant influence upon the pitch of the test signals. The importance of distortion products in this instance is not that they provide a simple cue to the pitch of a complex signal, but that they add further components to the amplitude spectrum which may, in certain circumstances, vary in amplitude as a function of the relative phases of the original components of a complex signal. If a three-component complex such as the SAM and QFM signals discussed above, with acoustic components at $f(\text{carr}) - g$, $f(\text{carr})$ and $f(\text{carr}) + g$, is subjected to a cubic nonlinearity, a series of seven distortion products are generated at the frequencies $f(\text{carr}) - 3g$, $f(\text{carr}) - 2g$, $f(\text{carr}) - g$, $f(\text{carr})$, $f(\text{carr}) + g$, $f(\text{carr}) + 2g$ and $f(\text{carr}) + 3g$. Since the phase of each distortion product depends upon the phases of the primaries, then the relative phases of the series of distortion products produced by the cubic transformation of a SAM and a QFM signal will also differ: The frequencies of some of these distortion products are equal to the frequencies of the primaries, so that the effective amplitude at each of the frequencies of the original components will depend upon the relative phases of the

original component and the distortion products of equal frequency. Thus the effective amplitude spectrum, that is the amplitude spectrum of the signal together with these distortion products, may change as a function of the relative phases of the components of the original signal. The effect such nonlinearity may have upon the amplitude spectra of the SAM and QFM signals employed by Ritsma and Engel and by Wightman is illustrated in figure 1.18, which shows the amplitude spectrum resulting from the addition of these products to the original signal. These spectra were generated by first calculating a 512-point digital representation of the acoustic signal of four periods duration, raising each point to the 3rd power, and then taking a weighted sum of this function and the original signal. Finally, a Fast-Fourier transform (Note 2) was applied to this sum to calculate the sine and cosine terms of the spectrum. The weights employed in the summing operation were those which were found to give a distortion product at $\underline{2f_2 - f_1}$ from a two-component signal whose level accorded approximately with that estimated by Smoorenburg (1972a), that is, at a level 17 dB below the level of the two equal-amplitude components. The cubic-nonlinearity is not the only plausible model of distortion products of this sort, and Smoorenburg (1972b, 1974) has suggested that an fractional exponent of about 0.6 provides a better account of the psychophysically observed distortion products. The effect of a nonlinearity with this exponent is similar to that of a cubic nonlinearity, except for the phases of the distortion products, and this nonlinearity will also give rise to phase-dependent differences in the amplitude spectra of distorted SAM and QFM signal. This argument is supported by the

Figure 1.18. The effect of a cubic nonlinearity on the amplitude spectra of SAM and QFM signals. Each signal has $f(\text{carr})$ at 11g Hz, and a modulation index of 2.55. Panel A shows the amplitude spectra of the original signals, panel B the spectra of their cubic distortion products, and panel C the sum of the two. The sum was obtained by the addition of the spectrum of the original signal $f(t)$ to the spectrum of $f(t)^3$, with weights of 1.0 and 0.00001 respectively. This weighting gives a distortion product at $2f_1 - f_2$ from a two-component signal which is 17 dB below the levels of the two primaries. All original components and distortion products for the SAM signal are in sine phase. For the QFM signal shown here, original components and distortion products at even-numbered harmonics are in sine phase, while those at odd-numbered harmonics are in cosine phase.



finding that the conditions under which relative phase effects upon pitch are observed coincide with those conditions under which these distortion products become significant (Buunen, and Bilsen, 1974; Buunen, Festen, and Bilsen, 1974).

If significant nonlinearities of this sort occur in the ear, then the cubic distortion products will affect not only the spectrum of the signal, but also its waveform, and thus also the temporal discharge pattern of the auditory nerve fibres. Hence, Licklider's temporal autocorrelation theory would also predict these relative phase effects, despite its essential insensitivity to phase differences. The spectral-pattern based theories can only predict the pitches of the signals used by Ritsma and Engel if the frequencies of two or more of the spectral components can be estimated from auditory frequency analysis. Because the distortion products will affect the waveform of an unresolved residue, this constraint is absent in the case of Licklider's theory. The role of distortion products may well prove to be of further importance in the development of theories of pitch perception.

The Pitch Shift and Auditory Nonlinearities.

A further illustration of the significance of auditory distortion products in the determination of residue pitch is provided by the experiments on the pitch shift of residue pitch. As was seen above, de Boer (1956) demonstrated that an arithmetic shift of the frequencies of the components of a complex signal results in an approximately proportional shift in

the residue pitch of the complex. Further studies of this pitch shift have revealed that the shift is greater than would be expected from the change in the intervals between peaks in the temporal fine-structure of the signal. For a two-component signal,

$$f(t) = \sum_n^{n+1} a_n \cos(2\pi f_n t),$$

the interpeak interval which a periodicity theory supposes to be responsible for residue pitch is equal to $1/\underline{f}$, and is equivalent to the time taken for $\underline{n}+1/2$ oscillations to occur at a frequency of $\underline{f}(\underline{n}+1/2)$ Hz. In the case of a signal whose spectral components are shifted arithmetically by the addition of a constant $\underline{\Delta f}$, Smoorenburg (1970) has shown that the shift in the major inter-peak interval is approximately $(\underline{n}+1/2)/\underline{\Delta f}$ sec, so that a periodicity theory would predict a pitch shift equivalent to $\underline{\Delta f}/(\underline{n}+1/2)$ Hz. For values of \underline{n} greater than 5, $\underline{f} = 200$ Hz, and signals presented at 40 dB above threshold, Smoorenburg found a pitch shift which exceeded $\underline{\Delta f}/(\underline{n}+1/2)$ Hz to a degree which increased monotonically with \underline{n} . These results are illustrated in figure 1.19. Even if the pitch of the signal was determined by the temporal structure of a filtered signal which was dominated by the lower component, then the maximum shift to be expected from the measurement of interpeak intervals would be equivalent to $\underline{\Delta f}/\underline{n}$ Hz, and this limit was also exceeded if \underline{n} was greater than 5.

Smoorenburg argued that if cubic or higher odd-order distortion products at frequencies $\underline{f}_3 = \underline{k}(\underline{f}_2 - \underline{f}_1)$, where $\underline{k} = 1, 2, 3, \dots$, were generated by the ear, then the presence of these additional components would effectively introduce

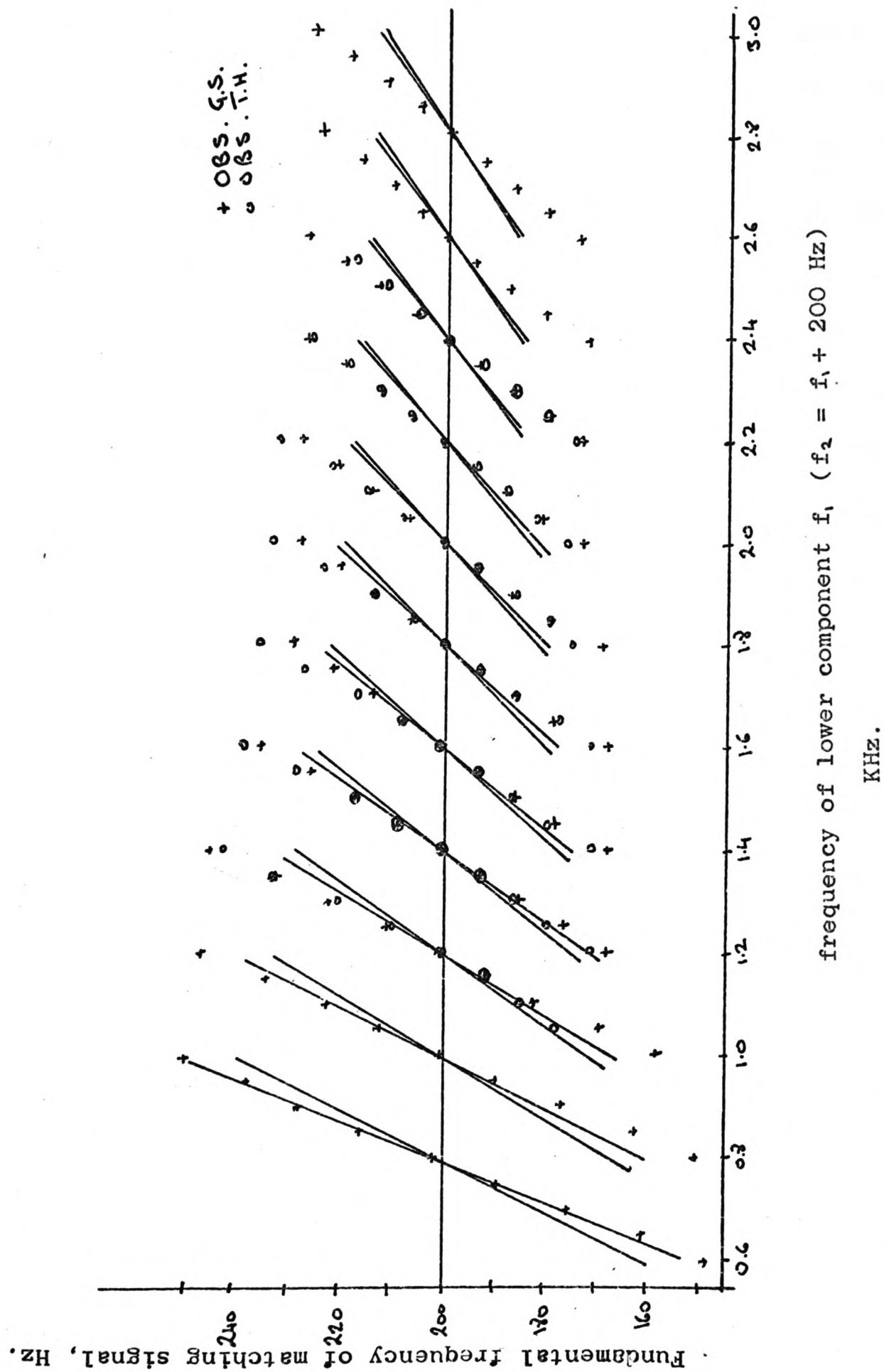


Figure 1.9. Matched pitches of two-component signals where $f_2 - f_1 = 200$ Hz. The solid lines represent the pitches expected if the pitch shift of the residue is $\Delta f/n$ or $\Delta f/(n+1)$. (Smoorenburg, 1970)

components representing harmonics below the nth, and that such additional components could well be responsible for the observed shifts exceeding the theoretical limits. Smoorenburg was able to demonstrate that such distortion products were indeed audible, and that the addition to the signals of low-pass filtered noise, which masked the distortion products at frequencies below the two original components, had the effect of reducing the observed pitch shifts to values closely corresponding to $\Delta f/n$ Hz. It is thus very likely that these distortion products were responsible for the discrepancy between the predictions and the data. The addition of the masking noise also resulted in the residue pitch of the complexes of high harmonic number (n greater than about 8) becoming less perceptible, which provides a further demonstration of the greater importance of information derived from the better resolved lower harmonics.

While the pitch shift of residue pitch has been discussed here with reference to quantitative predictions made on the basis of a periodicity theory, it should be emphasized that these results are also, after the same allowance is made for the influence of distortion products, consistent with the predictions of several spectral-pattern recognition models (Goldstein, 1973; Wightman, 1973b).

The Residue Pitch of Dichotic Two-component Signals.

One clear consequence of Schouten's periodicity theory is that monaural interaction of two or more components is a

necessary precondition for the perception of a residue pitch. Houtsma and Goldstein (1972) have presented results which suggest that melodies and musical intervals represented by the residue pitches of two-component complexes are identified as well with dichotically presented signals as with monotic presentation. If these results are accepted, then it follows that the preservation of the periodicity of the signal at the output of a monaural frequency analysis cannot be necessary for the perception of residue pitch, since in the case of a dichotically presented two-component complex, only one sinusoidal signal is present at each ear.

The two-component signals used by Houtsma and Goldstein were composed of two successive harmonics of a common fundamental frequency, each component being of equal amplitude. Two such complexes were presented successively, and the relation between the fundamental frequencies of the first and second signals was chosen to give one of the following musical intervals (the frequency ratios appear bracketed after each interval is named); an ascending third (4:5), an ascending minor third (5:6), an ascending second (8:9), an ascending minor second (15:16), or the same intervals in descending form. On any given trial, the observer was presented with two signals whose components were randomly chosen from the following possible pairs of harmonic numbers; $\underline{n}-1$ and \underline{n} , \underline{n} and $\underline{n}+1$, or $\underline{n}+1$ and $\underline{n}+2$. Results from two of the observers for both monotically and dichotically presented signals, at presentation levels of 20 and 40 dB above threshold, are shown in figure 1.20. At the higher presentation level, monotic performance is

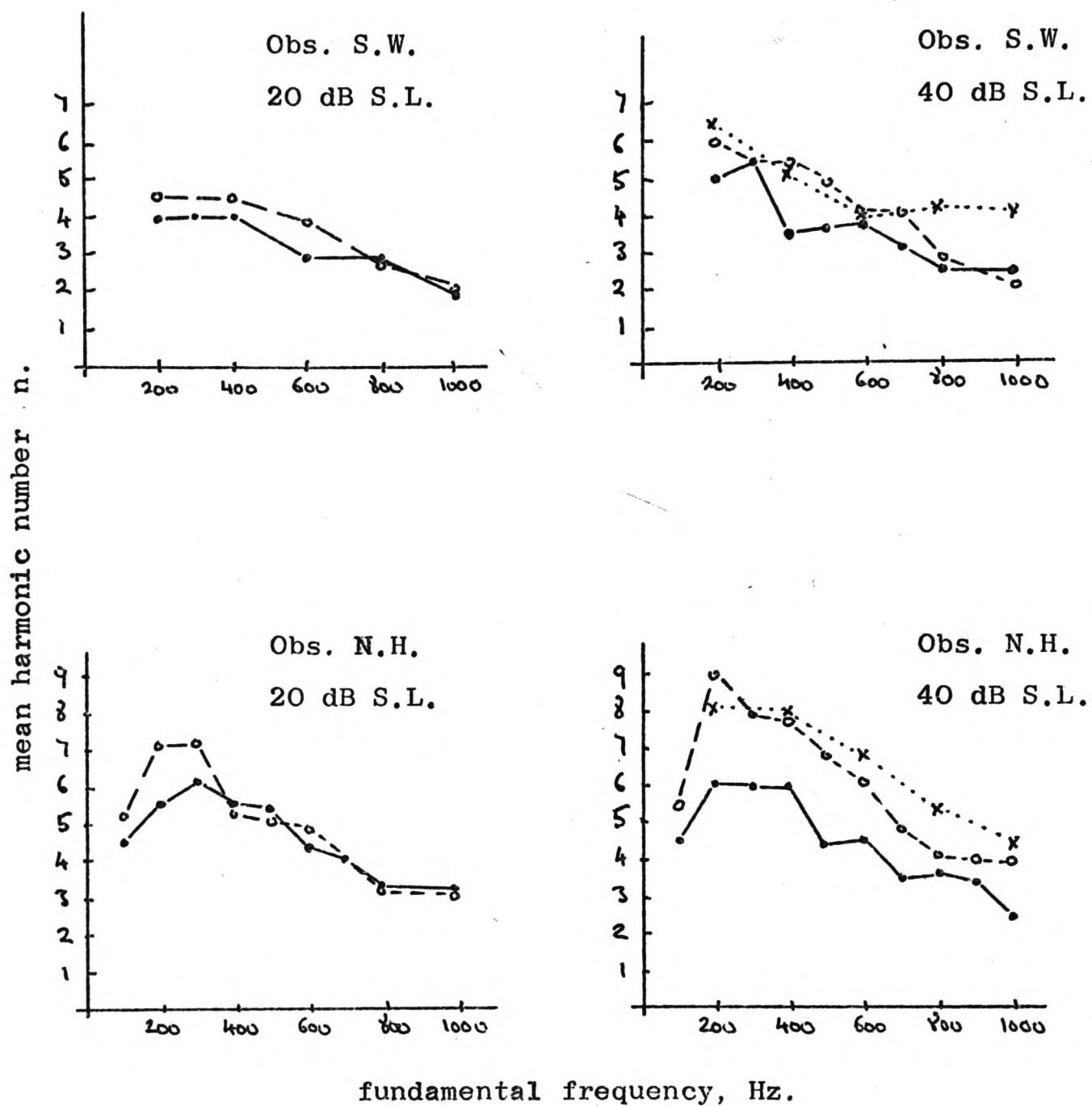


Figure 1.20. Values of n (approximate mean harmonic number) and fundamental frequency for which musical intervals represented by two-component complex tones were identified with an accuracy of 90%. The filled symbols represent dichotic presentation, the open ones monotic; the crosses in the figures for 40 dB S.L. represent performance when simulated distortion products were added to the dichotically presented signals. (Houtsma and Goldstein, 1972; figure after Plomp, 1976).

clearly superior to dichotic performance, but this superiority is not apparent at the lower presentation level. Houtsma and Goldstein assumed that the monotic superiority at the higher sound level was due to the influence of distortion products, which arise monaurally, and the addition of simulated distortion products to the dichotically presented signals did indeed lead to an improvement in performance (see figure 1.20); this result does not directly demonstrate that distortion products are responsible for the monotic superiority at the higher presentation level, but it is consistent with such an explanation of that difference.

The task of musical interval recognition does not provide a direct test of the assumption that the observers were responding on the basis of a residue pitch at the fundamental frequency of the signals, since other pitch cues, such as the pitches of the harmonics themselves, or residue pitches at frequencies other than the fundamental frequency might also account for some proportion of the correct responses. The possibility that the pitches of harmonics were employed as pitch cues is quite likely in this case, as all the observers were trained musicians. Houtsma and Goldstein, however, also report that one observer attempted unsuccessfully to identify musical intervals represented by pairs of two harmonics each presented non-simultaneously, so that the pitch cues from the pitches of the harmonics were still available. Taking this, and the random selection of harmonic number within each trial, into account, it may reasonably be concluded that the observers were responding on the basis of a residue pitch.

Subsequently, Hall and Peters (1980a, 1980b) have reported that a sequence of three non-simultaneous harmonics can result in the perception of a residue pitch under certain conditions. They found that with sufficiently brief presentation of the harmonics (40 msec each) separated by a very short inter-stimulus-interval of 10 msec, a residue pitch may be detected. This pitch is however only audible when the harmonics are presented just above their masked threshold in a background of noise; no residue pitch is heard with a silent background. An apparently related finding has been described by Houtgast (1976), who reported that a residue pitch was audible when only a single harmonic was presented. This effect occurred only where the harmonic was presented just above its masked threshold, and embedded in a sequence of signals having two or more components and an approximately constant residue pitch. It seems that these low signal-to-noise ratios somehow favour the perception of residue pitch, but the reasons for such an effect are obscure.

The results of Houtsma and Goldstein's study indicate that a residue pitch may be produced without any monaural interaction between the components of a complex signal, and therefore, it must be possible for pitch to be derived from a process which makes use of the central combination of dichotically presented information. Such a central combination may take place in either the frequency or the time domain, producing a central representation of the amplitude spectrum or of the periodicity of the original signal. The primary theoretical importance of these data is that they cannot be explained by Schouten's

periodicity theory, since that theory requires the monaural interaction of two or more components for the extraction of signal periodicity. Schouten's theory is also unable to explain Houtgast's (1976) report of a residue pitch from a single harmonic.

The Audibility of Residue Pitch and the Identifiability of Components.

Moore (1973b), has argued that if the spectral-pattern recognition view of the perception of residue pitch is correct, it follows that if none of the component frequencies of a complex signal are identifiable, no residue pitch can be extracted. Moore obtained psychophysical estimates of the minimum frequency separation between the frequencies of the components of a four-component complex at which the lowest component was identifiable, and compared these estimates with estimates of the minimum component spacing at which a residue pitch could be heard. For the first of these estimates, Moore employed a same-different task in which the observer heard in the first of two temporal intervals a complex tone having four components which were approximately equally spaced in frequency, and in the second interval, one of two comparison tones. On half of the trials, the comparison tone was a sinusoid at the same frequency as the lowest component of the complex signal, and on the remaining trials, the comparison tone was a sinusoid at a frequency \underline{d} Hz above that frequency. The difference \underline{d} was equal to $\underline{4DL}$, where \underline{DL} is the frequency difference limen at the frequency of the lowest component of the complex: Values for

the DL were taken from Shower and Biddulph (1931) for frequencies below 4000 Hz, and from Henning (1966) for higher frequencies. The values of d at the component frequencies used by Moore are shown in table 1.1. The complexes were inharmonic, and Moore reported that they produced no residue pitch. Each component was presented at a level of about 40 dB SL. The identifiability of the lowest component was investigated both with and without the addition to the complex of a band of noise, 1/3 of an octave wide, and centred 1/3 of an octave below that component: The level of this noise band was such that a similar noise band of equal level but centered on the lowest component would just mask that component. The observer's task was to judge whether the frequency of the comparison tone was the same as or different from that of the lowest component of the complex signal. From this task, Moore obtained an estimate of the probability of correct "same" and "different" responses as a function of the frequency separation between the component of the complex signal. The criterion adopted for a just-identifiable component frequency was a probability of correct "same" responses of 0.75, and the frequency separation yielding such performance was found by extrapolation. Moore reported an observed probability of incorrect "same" judgements of about 0.10. The critical frequency separations for the noise condition are shown in figure 1.21 as a function of the frequency of the lowest component of the complex.

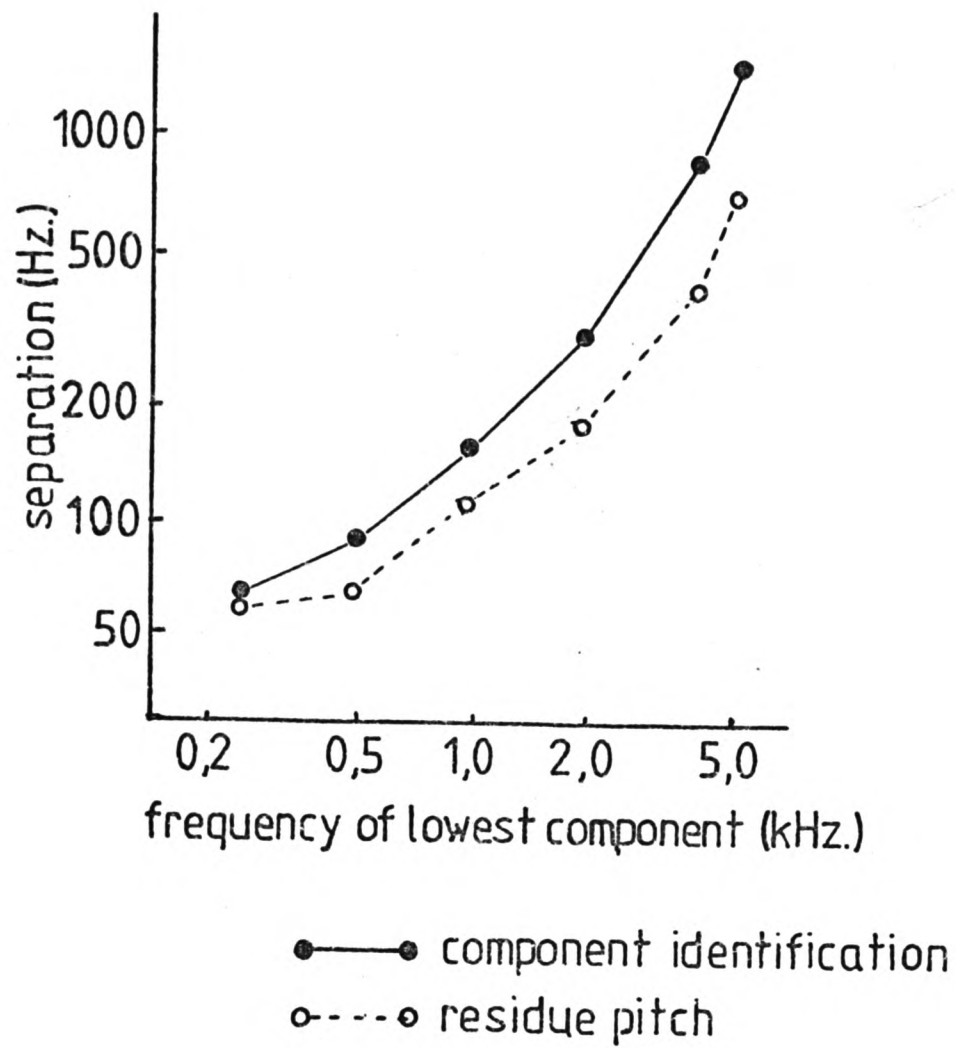


Figure 1.21 The minimum frequency separation between the components of a complex signal at which a) the lowest component is just identifiable, and b) a varying residue pitch can be heard, (after Moore, 1973b). Both sets of data were collected with a noise band 1/3rd octave wide centred 1/3rd octave below the lowest component of the complex in the case of the component identification data, and \underline{g} Hz. below the carrier frequency in the case of the residue pitch data.

Table 1.1. Approximate values of the frequency difference \bar{d} Hz between the two comparison tones for Moore's (1973a) investigation of the identifiability of the component frequencies of a complex signal.	
Frequency of lowest component, Hz	\bar{d} Hz
250	11
500	12
1000	16
2000	24
4000	64
5000	120

The minimum frequency separation between the components of a complex signal at which a residue pitch is audible was studied by a rather different method. The complex signal was produced by sinusoidal amplitude modulation at a frequency \underline{g} Hz of a carrier frequency $\underline{f(carr)}$ Hz whose frequency was modulated by about 2 Hz at a rate of about twice a second; considering only the two most prominent sidebands, the signal thus alternated between,

$$a) \ f(t) = a/(2m) \sin [2\pi(f(carr)-g)t] + a \sin (2\pi f(carr)t) + a/(2m) \sin [2\pi(f(carr)+g)t] \ ,$$

and,

$$b) \ f(t) = a/(2m) \sin [2\pi(f(carr)+2-g)t] + a \sin [2\pi(f(carr)+2)t] + a/(2m) \sin [2\pi(f(carr)+2+g)t] \ .$$

The modulation index \underline{m} was equal to 1, so that the two sidebands at $\underline{f(carr)-g}$ and $\underline{f(carr)+g}$ Hz were at a level 6 dB below that of the carrier frequency. The residue pitch produced by this signal corresponded to about \underline{g} Hz, and varied up and down about twice a second by about $2/(\underline{f(carr)}/g)$ Hz. This complex was presented at levels of 40 or 15 dB SL, and also at 15 dB SL with a noise band 1/3 octave wide centered 1/3 octave below $\underline{f(carr)}$ which would have masked the lower sideband, leaving only two

audible components. The minimum value of g Hz, which is equivalent to the frequency separation between the components, at which a varying residue pitch was audible in the noise condition is shown in figure 1.21 as a function of the carrier frequency $f(\text{carr})$. Also shown in the figure are Moore's estimates of the minimum frequency separation between components at which the lowest component of a four-component was identifiable. These are data from the condition with masking noise present, and a comparison of these results indicates that for component frequencies above 1000 Hz, a varying residue pitch was audible for frequency separations somewhat smaller than those at which the frequency of the lowest component of a four-component complex is identifiable. Moore concluded that this result was inconsistent with the spectral-pattern-recognition models of pitch perception, and demonstrated the importance of signal periodicity in determining residue pitch.

This conclusion is open to one serious criticism. It is not at all clear that Moore's criterion for the identifiability of component frequencies is equivalent to that for the audibility of a varying residue pitch: The latter criterion was purely subjective, and also open to influence by any other varying audible properties of the signal. If the performance of Moore's subjects is interpreted in terms of the Theory of Signal Detectability (Green and Swets, 1966; Macmillan, Kaplan, and Creelman, 1977), the value of the index of detectability d' for the component identification task, where the 'hit-rate' was 0.75 and the 'false alarm' rate was 0.10, is about 1.96.

For the detection of a modulated residue pitch in the second experiment, $\underline{d'}$ would be approximately 1, since this task is similar to Two-Alternative-Forced-Choice pitch discrimination, where for $\underline{d'}=1$, a pitch difference can be detected with a hit-rate of 0.76 and a false-alarm rate of 0.24. The value of $\underline{d'}$ for the component identification experiment strongly suggests that Moore has underestimated the identifiability of the frequency of the lowest component of the complex, and argues that his rejection of the spectral-pattern view cannot be sustained.

The Pitch of Modulated and Delayed Noise.

This review has up to now been restricted to pitches resulting from tonal signals, that is, signals containing periodic components. There have also been a number of reports of pitch or pitch-like sensations arising from signals which have been thought not to contain periodic components; such signals have been constructed from aperiodic noise. Several studies (Harris, 1963; Miller, and Taylor, 1948; Pollack, 1969) have demonstrated that periodically interrupted white noise, with interruption rates of between about 20 and 2000 Hz, gives rise to a pitch-like sensation which varies with the interruption rate. The observers in these studies were able to adjust the frequency of a pure-tone to match this pitch, and, disregarding errors of matching of an octave, the accuracy of these matches was between about 2% and 5%. Several studies of the pitch of amplitude-modulated white noise have also been reported, (e.g. Burns and Viemeister, 1976; Patterson, and

Johnson-Davies, 1977), and have shown that these signals also exhibit a weak pitch corresponding to the modulation rate. Burns and Viemeister were able to demonstrate that the pitch of sinusoidally amplitude modulated white noise could support the identification of melodies and musical intervals, which suggests that this perceptual aspect of modulated noise does have the properties of a true pitch. The accuracy of Burns and Viemeister's observers in these tasks was, however, considerably inferior to their accuracy when the melodies or musical intervals were constructed from sinusoidal signals. As Houtsma and Goldstein (1972) have shown, recognition for melodies and musical intervals constructed from the residue pitch of two-component harmonic signals is of comparable accuracy to recognition with sinusoidal signals. The pitch of interrupted or modulated noise is thus considerably weaker or more ambiguous than the pitch derived from tonal signals.

The long-term power spectrum of these symmetrically modulated signals is flat, and Burns and Viemeister claimed that their results provided conclusive evidence of the importance of the detection of signal periodicity in pitch perception. However, it has now been shown (Pierce, Lipes, and Cheetham, 1977), that the short-term spectra of these signals are not flat, but contain certain regularities at frequencies related to the interruption or modulation rate, so that it is not necessarily the case that signal periodicity is responsible for this pitch.

If a white noise signal is added to a copy of itself

delayed by some time interval τ sec, the spectrum of the sum shows peaks at frequencies which are multiples of $1/\tau$ Hz. Bilsen and Ritsma (1970) have reported that such a signal has a pitch corresponding to $1/\tau$, which remains audible when the signal is high-pass filtered above $1/\tau$ Hz. This pitch is thus similar to the residue pitch of tonal signals. Bilsen and Ritsma have also shown that the pitch of such noise signals is similar to residue pitch in other respects: The salience of this pitch is greatest when energy in the region of $4/\tau$ Hz is present, which is consistent with the dominance region for the pitch of tonal signals (Ritsma, 1967), and for a delay period of about 5 msec, which corresponds to maximal extent of the existence region for fundamental frequencies of about 200 Hz (Ritsma, 1964).

An intriguing result has been described by Bilsen and Goldstein (1974), who report that white noise, presented directly to one ear, and to the other ear with a time delay, can give rise to a pitch corresponding to the reciprocal of the delay time. With this signal, the peripheral signals at each ear show neither any spectral variations nor any periodicity related to the perceived pitch, so that the pitch of the signal, whether coded by periodicity or spectral cues, must be derived centrally. It has been shown (Bilsen and Goldstein, 1974; Bilsen, 1977) that the appropriate central spectral-pattern can be generated by a system such as Durlach's (1972) Equalization and Cancellation model of binaural detection.

According to the investigators who have studied the pitch

of noise signals, this pitch is both less salient and less precise than that of tonal signals, and its existence provides no compelling reasons for a modification of any of the conclusions reached above.

Summary.

The studies reviewed here provide a conflicting body of evidence. If residue pitch is derived from the periodicity of the signal as preserved peripherally as the result of the interaction of several components, then several phenomena, particularly the dominance of low harmonics (Plomp, 1967; Ritsma, 1967) and the residue pitch of dichotically presented two-component complexes (Houtsma and Goldstein 1972), are difficult to explain, while a theory based upon the use of spectral cues has particular difficulty dealing with the effects on pitch of the relative phases of components (Moore, 1977; Ritsma, and Engel, 1964). A spectral theory also has difficulty with the extension of the existence region to complexes consisting of components which are, if the critical bandwidth is an appropriate estimate of resolving power, too close in frequency to be resolved. Since a theory based upon the recombination of the temporal information provided by individual components is also subject to the limits imposed by auditory frequency analysis, and additionally by limited temporal accuracy, then the difficulties here are similar to those suffered by spectral-pattern recognition theories.

The most prominent feature of the studies reviewed here is

that the better resolved components of complex signals play a dominant role in the determination of residue pitch, which clearly supports the spectral-pattern-recognition theories rather than Schouten's periodicity theory, whose predictions are at variance with the great majority of the evidence. The possibility that both spectral pattern and periodicity are necessary for a complete theoretical treatment cannot be rejected, so that a theory such as Licklider's neural autocorrelation model, which is able to utilize both of these sources of information, may yet provide the most satisfactory account of pitch phenomena.

CHAPTER 2

Abstract.

This chapter discusses a number of theories of the perception of residue pitch based upon spectral-pattern recognition. The models proposed by Goldstein (1973) and Wightman (1973b), are discussed in detail. A general property of all of these theories is that the predicted residue pitch can be considered as a random variable derived more or less directly from a weighted sum of a series of random variables representing internal estimates of the frequency or the period of each of the several components of a complex signal. In contrast, the residue pitch predicted by Schouten's periodicity theory and by an unmodified version of Licklider's neural autocorrelation theory can be considered to be derived from a single random variable, this being an internal estimate of the period of an unanalysed residue. This distinction, together with the assumption that the components of a complex signal are well-resolved, which implies that the internal representations of the components can be considered to be statistically independent, leads to a psychophysically testable difference between the predictions of the spectral-pattern recognition models and those of the periodicity model. According to the

spectral-pattern recognition models, $\sigma(f'_0)/f_0$, the relative standard deviation of the theoretical distribution of the residue pitch in the region of the true fundamental frequency of a harmonic signal, should be smaller than the minimum of $\sigma(f'_j)/f_j$, the relative standard deviations of the theoretical internal distributions of estimates of the component frequencies. According to the periodicity model the reverse prediction is made.

Predictions are also developed from the spectral-pattern recognition models for the pitch pdf of two two-component signals, which are the signals employed in the experiments described in the subsequent chapters. Several of the experiments described below are intended to produce psychophysical estimates of the parameters $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ under a variety of presentation conditions. If the latter parameter is known for each component of the signal, a computer simulation of these models can predict the areas of each mode of the multimodal pitch pdf. Several experiments, described in Chapter 10, were designed to compare these predictions with psychophysical estimates of these areas.

Introduction.

Spectral-pattern recognition theories of residue pitch perception all have the common purpose of deriving an estimate of the period or fundamental frequency from the information provided by a frequency analysis of a complex signal. This information can include contributions from each ear, so that all such models can extract a residue pitch from two dichotically presented components. The notion of a template representing generally the spectrum of a periodic signal, and which may be compared with the spectrum of the presented signal was introduced earlier (Chapter 1, pp 20-21), and such a matching process is analogous to, although not necessarily identical to, the processes represented by all models of this class.

A notable strength of the spectral-pattern recognition models of pitch perception described by Goldstein (1973) and Wightman (1973b) is that these models are described in sufficient formal detail that a theoretical probability density function (pdf) for the pitch of a given auditory signal may be estimated from a computer simulation. The first part of this chapter contains a formal description of these two models. The second part of the chapter is concerned with some predictions of these and other models, and with the computer simulation methods by which the theoretical pitch pdfs are generated.

Goldstein's Optimum-Processor Model.

The model described by Goldstein (1973), and qualified by Gerson and Goldstein (1978) and Goldstein, Gerson, Srulovicz and Furst (1978), has its mathematical basis in the maximum-likelihood estimation of integer ratios relating the frequencies of the components of a presented signal. If a well-fitting set of ratios may be found, and if the signal is periodic and precisely coded, then the integers will be equivalent to the harmonic numbers of the components of the signal. From these harmonic numbers and the values of the component frequencies, an estimate of the fundamental frequency of the signal may be made. If a number of estimates of these harmonic numbers are made, then each set of estimates may be evaluated for its goodness-of-fit with the ratios between the component frequencies. This process may also be applied to inharmonic signals, which have no true fundamental frequency.

Goldstein's (1973) formulation of his 'optimum-processor' model starts with the following postulates:-

1/ All stimuli are assumed to be periodic, so that the pitch extraction process attempts to estimate the fundamental frequency of any signal, even if no true fundamental frequency exists.

2/ The process only employs information given by internal estimates of the frequencies of the components of the signal, the phase and amplitude of each component being ignored. It

follows from this constraint that only resolved components can contribute to the process. The manner in which frequency is coded is unspecified, and the model does not exclude the possibility of a temporal code.

3/ For the purposes of estimating harmonic numbers, the signal is assumed to consist of a series of successive harmonics of a common fundamental frequency. In a subsequent revision of the model (Gerson and Goldstein, 1978), this constraint is relaxed, and it is merely assumed that the process regards signals consisting of successive harmonics as a priori more likely, and that the estimation process initially attempts to find a set of successive harmonic numbers which match the signal, and only evaluates other possibilities if no such solution is found, or if the a priori likelihood of other types of signal is modified by immediate past experience.

4/ The internal estimates of component frequencies are assumed to be noisy, and distributed as Gaussian random variables, the representation f'_j of the j th component having a population mean $\mu(f'_j)$ approximately equal to the true component frequency f_j , and a standard deviation $\sigma(f'_j)$ which is assumed by Goldstein to be dependent only upon frequency, so that $\sigma(f'_j) = \sigma \cdot f_j$.

The model thus requires only the information that a given number of simple components are present in the signal, or are resolved by auditory frequency analysis, and the estimates of the frequencies of these components.

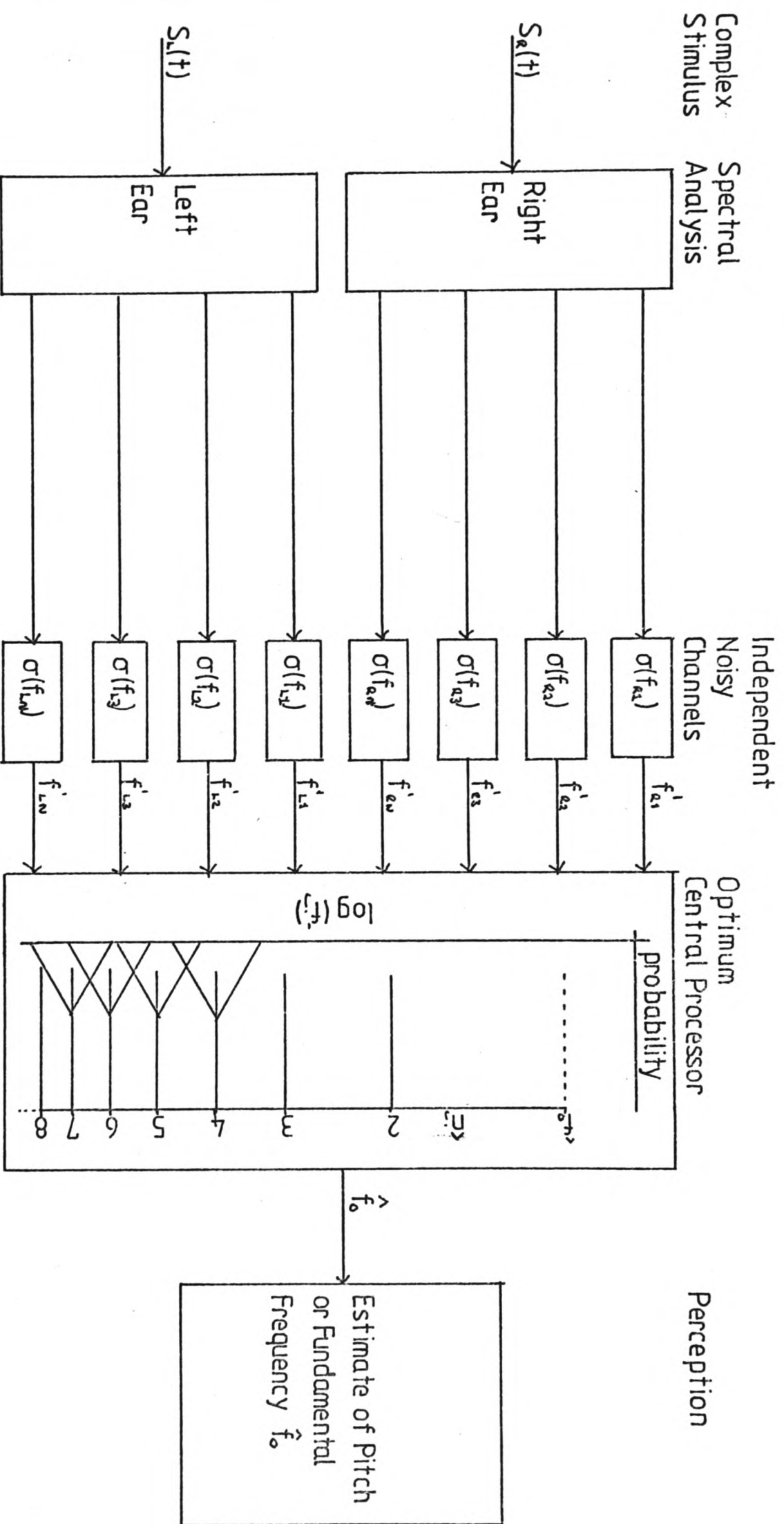


Figure 2.1 Schematic representation of Goldstein's (1973) Optimum Processor model. Estimated component frequencies are conveyed to the central pitch processor by a series of noisy channels.

For a signal containing k resolved components, the jth component is represented by an estimate of its frequency \hat{x}_j , that is, a sample from the population distribution f'_j , and an associated estimated standard deviation $\sigma(\hat{x}_j)$. The fit of an assumed set of estimated harmonic numbers \hat{n}_j , may be found according to;

$$\epsilon^2 = \sum_{j=1}^k \hat{x}_j^2 / \sigma^2(\hat{x}_j) - [(\sum_{j=1}^k \hat{n}_j \hat{x}_j / \sigma^2(\hat{x}_j))^2 / \sum_{j=1}^k \hat{n}_j^2 / \sigma^2(\hat{x}_j)] \quad (2.1)$$

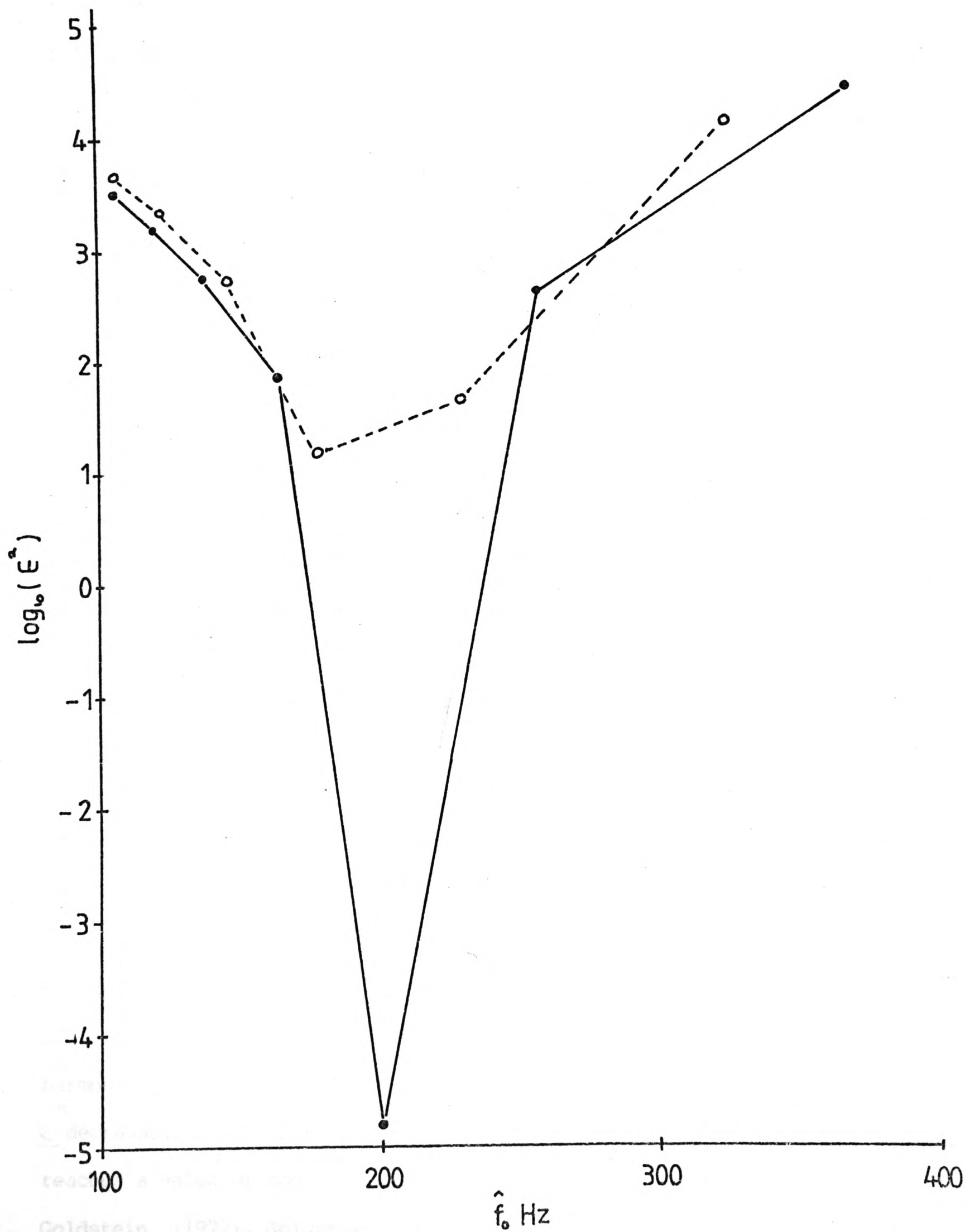
The pitch or estimated fundamental frequency \hat{f}_0 for a given set of estimated harmonic numbers is then;

$$\hat{f}_0 = \sum_{j=1}^k [\hat{n}_j \hat{x}_j / \sigma^2(\hat{x}_j)] / \sum_{j=1}^k \hat{n}_j^2 / \sigma^2(\hat{x}_j) \quad (2.2)$$

and is thus a weighted sum of k random variables. The best-fitting estimate of fundamental frequency is found by attempting to minimize the squared error term ϵ^2 . The stages of the model are shown schematically in figure 2.1. If all plausible sets of integer estimates of \hat{n}_j are used in the search for minimum error, then this process is equivalent to searching for that location on a frequency dimension at which a template representing a generalised harmonic spectrum is most highly correlated with the estimated frequencies \hat{f}'_j of the signal, and the estimated pitch of the complex signal is then equivalent to the frequency of the fundamental component of the template. Figure 2.2 shows the values of the error term for a two component signal where $\sigma(f'_j)/f_j = 0.01$ against a series of estimated fundamental frequencies; here the signal was assumed to consist of two successive harmonics.

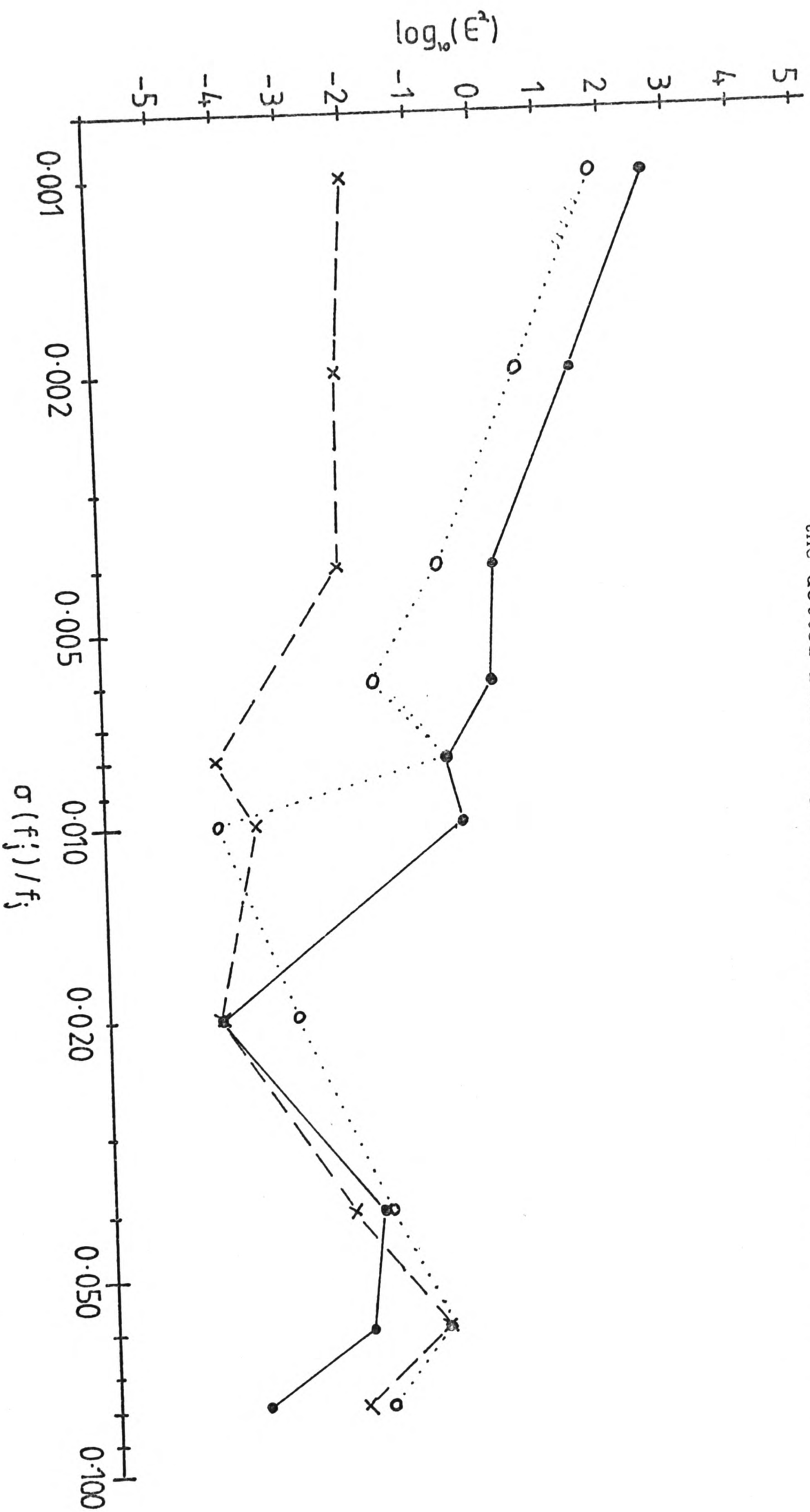
This model requires an estimate of $\sigma(f'_j)$ in order to estimate the pitch of a complex signal, and Goldstein has assumed that $\sigma(f'_j)$ may be derived from the sample estimate

Figure 2.2 The error term ϵ^2 for Goldstein's model as a function of the estimated fundamental frequency \hat{f}_0 . Values shown are for a harmonic two-component signal (800 and 1000 Hz, solid line) and an inharmonic complex containing components at 700 and 900 Hz (dashed line).



\hat{x}_j , by the equality $\underline{\sigma(f'_j)} \approx \underline{\sigma \cdot \hat{x}_j}$. In practice, $\underline{\sigma(f'_j)}$ may also depend upon other factors, particularly the signal-to-noise ratio and the duration of the signal (see Chapter 3), so that this assumption presents some difficulties. It is, however, plausible that $\underline{\sigma(f'_j)}$ may be estimated from an amplitude or power spectrum derived from an auditory frequency analysis, since the necessary information would be provided by the height-to-width ratio of each spectral peak. Goldstein's discussion of the model suggests that the dominant component of $\underline{\sigma(f'_j)}$ is internal noise, and that this noise is present in a set of neural channels having the specific purpose of conveying component frequency estimates to the pitch extractor. In this model, the assumption of a noisy representation of component frequencies is not merely a recognition of the limitations of auditory coding, but is also functional, since a signal containing very precisely coded inharmonic components could not be assigned any pitch value with a small associated error term . For such a signal, the contribution of the internal representation of each component to the error term will be large, and if this term is weighted by the reciprocal of a small $\underline{\sigma^2(f'_j)}$, the total error for any estimated pitch will necessarily be large. The dependence of the error term upon the variability of component frequency representation for harmonic and inharmonic two-component signals is shown in figure 2.3, from which it may be seen that while the error term for a harmonic signal is consistently small, for an inharmonic signal, $\underline{\epsilon^2}$ decreases with increasing $\underline{\sigma(f'_j)}/f_j$ until that parameter reaches a value of about 0.010. From the data of Houtsma and Goldstein (1972), Goldstein (1973) has estimated that the noise

Figure 2.3 The error term for Goldstein's model as a function of $\sigma(f'_j)/f_j$. Values are plotted for three signals, and in each case, the value plotted is the smallest for that signal. The dashed line is for a harmonic two-component signal containing components at 800 and 1000 Hz. The solid line is for a signal containing components at 700 and 900 Hz, and the dotted line a signal containing components at 900 and 1100 Hz.



in the channels conveying component frequency estimates to the pitch extractor is such that $\sigma(f'_j)/f_j$ is approximately equal to 0.01. If this model is considered as a formalisation of the template matching process described above (Chapter 1, pp 20-21), the effect of Goldstein's assumption that noise is deliberately added to the component frequency estimates is equivalent to the effect of the assumption that the template itself is noisily represented; this latter assumption would be indistinguishable from the assumption that estimated harmonic numbers are not integers, but real random variables.

The probability density functions predicted by Goldstein's model for the pitch of two two-component signals are evaluated below; since an analytic solution of the model is not available, the necessary calculations will be made by computer simulation, for which the procedure is detailed below.

Terhardt's Multiple Cue Theory.

A model similar to that described by Goldstein was proposed by Terhardt (1974), in which it is assumed that each resolved component of a complex signal provides a series of pitch cues, these having values equal to each integer sub-multiple, or subharmonic, of the frequency of the component. If the signal is periodic, then the n th subharmonic of the n th harmonic will be equal to the fundamental frequency, and Terhardt argues that it is the coincidence of such pitch cues from several components at one frequency which provides the basis for the selection of a pitch or fundamental frequency estimate. He also suggests,

following Thurlow (1963), that the ability to utilize the coincidence of pitch cues is learned.

Walliser (1969a) has suggested a similar process, but includes also the assumption that the envelope rate of the signal is used as a cue to the frequency region in which a pitch is to be sought. According to this model, only a single component frequency estimate is necessary, since the residue pitch may be assigned the value of the subharmonic of this component which is closest to an estimate of the envelope rate.

Both Terhardt's and Walliser's models lack a systematic treatment of the variability of component frequency and fundamental frequency estimates. These models are otherwise similar to Goldstein's model, and will receive no further attention .

Wightman's Autocorrelation model.

Wightman's (1973b) autocorrelation model represents, superficially at least, a markedly different approach to the problem. While Goldstein's model operates entirely in the frequency domain, Wightman's model ultimately produces a temporal representation of the signal from which the period is estimated. A further difference between these two models is that Wightman's model also takes account of the relative amplitudes of the components of a complex signal. The model assumes, as does Goldstein's, that the pitch extraction process begins with the derivation, from an auditory frequency analysis

of limited resolution, of an estimate of the spectrum of the signal, which is derived from the sum of a series of estimates of the frequency of each resolved component. For mathematical convenience, Wightman assumes that the estimated spectrum is an approximation to the power spectrum of the signal. The autocorrelation of the power spectrum is then calculated by some neural process approximating an inverse Fourier transform, and from the autocorrelation, an estimate of the period, the reciprocal of the fundamental frequency, is made.

These processes are illustrated by Wightman as follows. Consider an idealized and truly periodic signal, consisting of k components, each of unit amplitude, but with arbitrary relative phases,

$$f(t) = \sum_{j=1}^k (\cos w_j t + \theta), \quad (2.3)$$

where w_j is the angular frequency of the j th component in radians/sec., and θ is an arbitrary value representing the initial phase angle of the j th component. The power spectrum, that is the square of the amplitude spectrum, will be,

$$|F(w)|^2 = \sum_{j=1}^k \pi^2 \delta(w - w_j), \quad (2.4)$$

where $\delta(x) = 0$ if $x \neq 0$ and 1 otherwise. The autocorrelation function is then found by taking the inverse Fourier transform of the power spectrum;

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{j=1}^k \pi^2 \delta(w - w_j) \cos w \tau dw, \quad (2.5)$$

which may be simplified to:-

$$R(\tau) = \pi/2 \sum_{j=1}^k \cos w_j \tau, \quad (2.6)$$

and may be seen to be equivalent to the signal apart from the constant factor $\pi/2$, and the constant value of 0 for θ . A schematic representation of the stages of the model is shown in

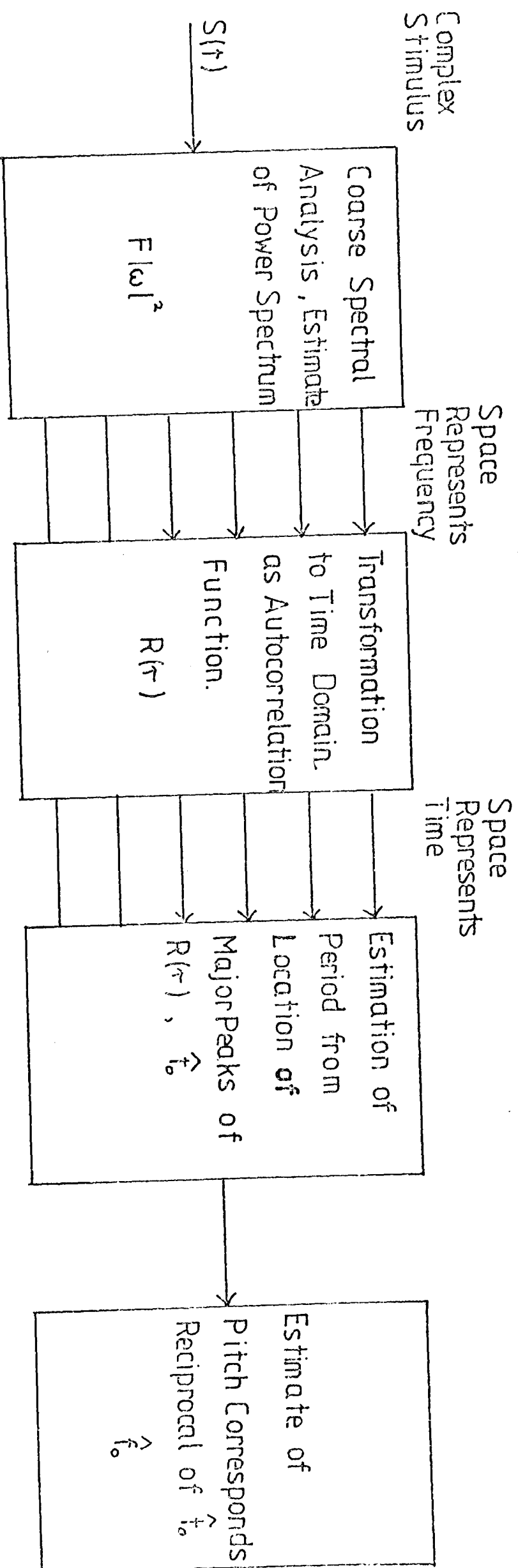


Figure 2.4 Schematic representation of Wightman's (1973) Autocorrelation model.

figure 2.4. The effect of this series of transformations is thus to produce a function in the time-domain which differs from $f(t)$ primarily in that the relative phase information is lost. If the components of the signal $f(t)$ are not of equal amplitude, then the power spectrum will emphasize the components of higher amplitude, and this emphasis will remain in the resulting autocorrelation function. The function may be normalised by division by $R(0)$, resulting in a function with maximum and minimum values of +1 and -1 respectively. The autocorrelation function may also be considered as the correlation of the waveform of the signal with the same waveform delayed in time by τ , with τ valued from $-\infty$ to $+\infty$, or some reasonable approximation thereof, so that the autocorrelation function can also be calculated entirely within the time domain, and the resulting function will be identical to the function calculated by inverse Fourier transformation of the power spectrum.

Since the computation of an autocorrelation function involves the calculation of a power spectrum and an inverse Fourier transform, and these calculations may be carried out for aperiodic signals of limited duration, the application of Wightman's model to such signals presents no formal difficulty. With such aperiodic signals, however, the autocorrelation function becomes decreasingly similar to the waveform of the signal. As the amplitude spectrum of the signal becomes increasingly unlike a discrete line spectrum, that is, as the proportion of random noise in the signal increases, or the signal duration decreases, the autocorrelation function becomes

progressively damped over time. The peaks of the envelope of the function show progressively smaller maxima; for a signal of finite duration, the envelope peaks reach their smallest values at a value of τ equal to half of the duration of the signal, since the resulting autocorrelation function is truly periodic over a period equal to the duration.

If the signal is periodic, and provided that the damping of the function is not excessive, then the autocorrelation function shows its largest peak where τ is equal to the period t_0 (neglecting the peaks at $\tau = 0$). Wightman's model assumes that the autocorrelation function derived from the presented signal is scanned for major peaks, and that the estimated period is taken as the value of τ at the greatest of these peaks. The model further proposes that the amplitude of each peak determines the 'strength' or salience of the corresponding pitch, a stronger pitch having a greater probability of being assigned to the signal, although the relationship between this probability and the value the peak is not defined in the model.

As with Goldstein's model, the predictions of the autocorrelation model will be evaluated by computer simulation. In carrying out these simulations, the noisiness of the power spectrum is, as in Goldstein's model, a free parameter, and while Wightman discusses such variability in terms of the equivalent bandwidth of auditory frequency analysis, here this noise process will be treated as the standard deviation of a Gaussian probability density function for the estimated value of each component frequency, this distribution being equivalent to

the distribution $\underline{f'_j}$ in the above discussion of Goldstein's model.

The claim was made above that Wightman's model was an example of the spectral-pattern recognition theories, but this is not immediately apparent from the above formulation. The similarity arises from the direct reciprocity of time and frequency, whereby the autocorrelation at a given time delay \underline{T} of some signal is analogous to the correlation between a spectral template having the fundamental frequency $\underline{1/\tau}$ and the spectrum of the signal. The process of harmonic number estimation in Goldstein's model is analogous to the selection in Wightman's model of a delay interval about which a peak is sought.

Thus far, it is not clear to what extent the predictions of these two models differ, and de Boer (1977) has argued that the two models have a number of points of equivalence. The computer simulations of the two models with two-component harmonic signals described below should provide some evidence bearing upon this question. One basic difference between the two models is that Wightman's model, unlike Goldstein's, is inherently sensitive to the relative amplitudes of resolved components. It is of interest that Ritsma (1967) has reported that harmonics between the 3rd and 6th dominate the perceived residue pitch even when higher components are at levels up to 40 dB above the lower components, which argues against a direct sensitivity to the relative amplitudes of resolved components.

Component Frequency Estimate Variability and Residue Pitch.

According to the spectral-pattern recognition models as formulated above, if the relative standard deviation of the internal representation of frequency $\sigma(f'_j)/f_j$ is systematically varied, then the parameters of the pdf of the predicted pitch values change in consequence. It will be shown below that an increase in $\sigma(f'_j)/f_j$ has two major consequences: The pitch pdf becomes increasingly multimodal, and the areas of the modes above and below the mode at the true fundamental frequency increase, so that, in the case of a harmonic signal, the salience of pitches at frequencies other than the true fundamental, and the probabilities of the assignment of such pitches increase. The multimodality of the pitch pdf becomes particularly apparent for signals containing only a small number of components. In terms of psychophysical observations, these changes in the pdf would correspond to a decrease in the probability that an observer will match the residue pitch of a harmonic signal at its fundamental frequency. The standard deviation of each mode of the pdf also increases, so that estimates of the residue pitch within each mode become more variable. The psychophysical correlate of this latter change would be an increase in the pitch difference limen, that is, the pitch difference between two residue pitches which is just detectable. If Goldstein is correct in his (1973) assumption that $\sigma(f'_j)/f_j$ is limited by an internal noise specific to the pitch extraction process, then there is no obvious procedure for the direct psychophysical estimation of $\sigma(f'_j)/f_j$. If, however, $\sigma(f'_j)/f_j$ is largely or wholly composed of noise due

to external sources and to peripheral stimulus encoding, $\sigma(f'_j)/f_j$ can be estimated from a frequency discrimination task provided that a task may be designed such that the observer is presented with all of the signal components, and remains able to listen for frequency differences in just one component. Such a task is described in Chapter 5.

The Variability of an Estimated Residue Pitch at the Fundamental Frequency.

Predictions of the spectral-pattern recognition models.

Goldstein (1973) has shown that, for his model, the relationship between $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$, where $\sigma(f'_0)/f_0$ is the relative standard deviation of the mode of the pitch pdf having the mean f'_0 which is approximately equal to the true fundamental frequency of the signal f_0 , can be approximated by

$$1/[\sigma(f'_0)/f_0]^2 = \sum_{j=1}^K 1/[\sigma(f'_j)/f_j]^2 \quad (2.7)$$

for a two-component signal. Considering the expression given above as equation 2.2 as the weighted sum of a series of weighted independent random variables, equation 2.7 follows from the appropriate expression for the variance of such a sum (Hays, 1963), where

$$\sigma^2(f'_0) = \sum_{j=1}^K \{ \sigma^2(\hat{x}_j) \cdot [\hat{n}_j / \sigma^2(\hat{x}_j)]^2 \} / \left[\sum_{j=1}^K \hat{n}_j^2 / \sigma^2(\hat{x}_j) \right]^2 \quad (2.8)$$

In the case of Wightman's autocorrelation model, no derivation of the relationship between these standard deviations has been described. Conventional treatments of the distribution of the autocorrelation function (Papoulis, 1965) are generally

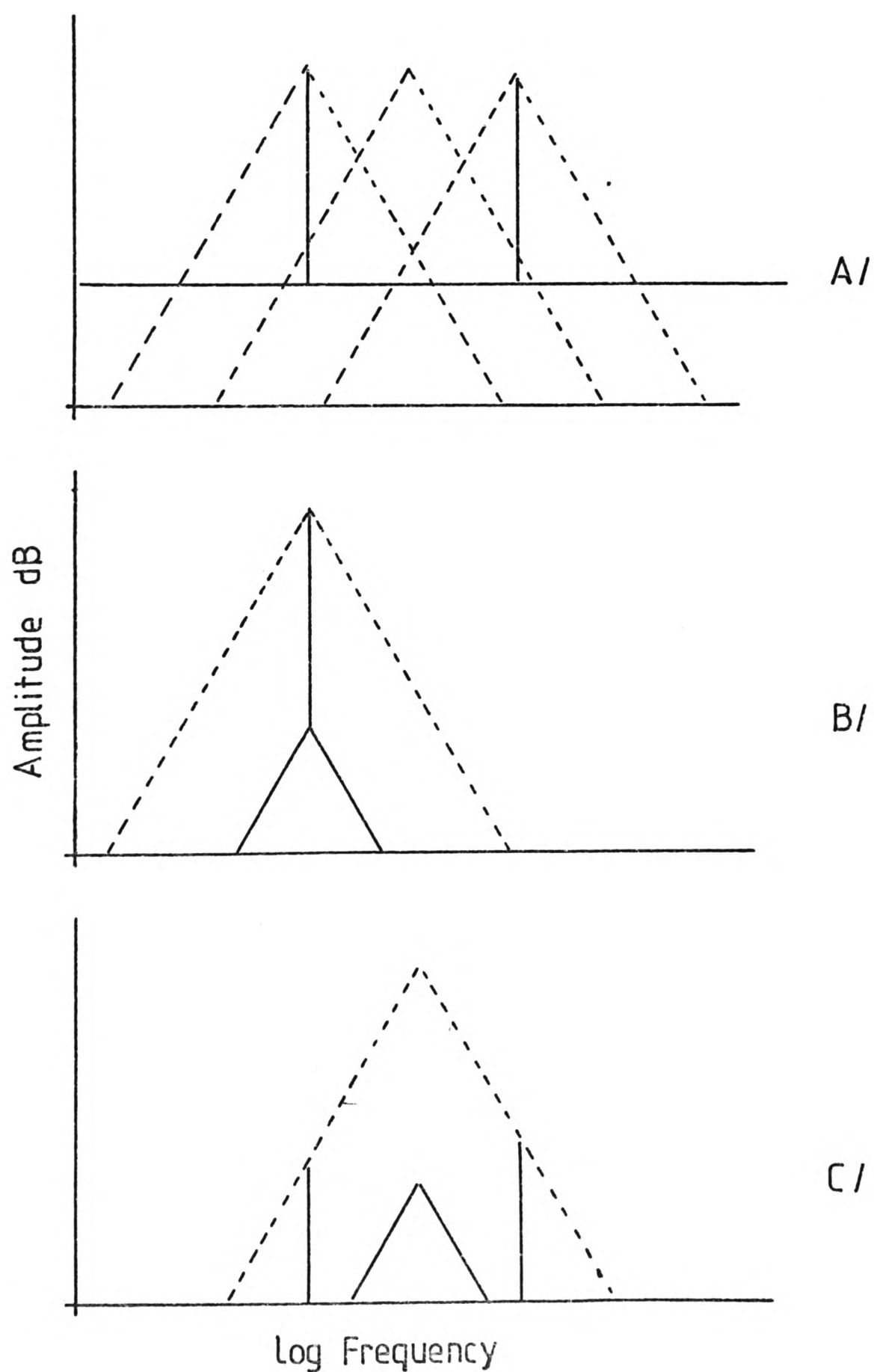
concerned with the probability that the value of the function at a given point in time is greater or less than some limit, whereas the statistic required here is the variance of the distribution of time values at which the function shows a peak. Since the value of the autocorrelation at any given delay time will be a random variable, and successive samples of the function are correlated, this variance is not easily calculated. Here, this relationship has been estimated by the simulation methods described below. If $\sigma(f'_j)$ is taken as the standard deviation of a Gaussian pdf which is assumed to be equivalent to an internal estimate of the power spectrum of the component f_j , the relationship between $\sigma(f'_0)$ and $\sigma(f'_j)$ is found to be the same as that derived by Goldstein for the optimum-processor model (this result is described more fully below). This result is intuitively plausible, as the estimated power spectrum which is transformed to yield the autocorrelation function is a sum of a series of independent random variables, each of these being an internal estimate of the frequency of one of the resolved components of the signal, and thus equivalent to the sum of the distributions f'_j in Goldstein's model.

The extension of Licklider's (1959) autocorrelation theory discussed in Chapter 1, which allows the use of temporally coded estimates of the frequencies of resolved components, makes this theory generally similar to Wightman's model where only well-resolved components are present, and the predictions of the two theories should be very similar.

The predicted value of $\sigma(f'_j)/f_j$ from Schouten's Periodicity Theory

While Schouten's (1940) periodicity theory is not favoured by the review presented in the previous chapter, it is nevertheless worthwhile to attempt a description of the predictions which this model might make for the relations between $\sigma(f'_j)/f_j$ and the residue pitch pdf. With this model, independent estimates of individual component frequencies play no part in the determination of pitch so that there need not be a direct relation between such estimates and the pitch pdf. As has been argued above, where a complex signal has two components which are sufficiently different in frequency to be well resolved by auditory frequency analysis, the best estimate of the frequency of a single component will be derived from a channel in which other components have a minimal effect and in which the best signal-to-noise ratio for that component is found, since the signal-to-noise ratio will impose limits to the precision of any frequency code. The periodicity of the signal however, will be best represented in a channel responding equally to both components, that is a channel corresponding to a filter having a centre frequency between the two component frequencies. In such a filter, the signal-to-noise ratio will be worse than that in the two filters centred on the two component frequencies by a factor determined by both the slopes of the skirts of this filter, and by the difference between its centre frequency and the component frequencies; for well resolved components, this factor will be fairly large. The information available from such a set of filters is illustrated

Figure 2.5. The coding of periodicity and component frequency for two well-resolved components. Panel A/ shows the signal and noise spectrum prior to filtering, while B/ and C/ illustrate the optimal positioning of the auditory filters for the coding of the frequency of the lower component, and the periodicity of the waveform. The signal-to-noise ratio for the periodicity is smaller than that for component frequency.



in figure 2.5. From this reasoning, it may be concluded that, for well-resolved components, $\sigma(f'_0)/f_0$ will be greater than $\sigma(f'_j)/f_j$, in contrast to the prediction of the spectral-pattern-recognition models, where $\sigma(f'_0)/f_0$ should be smaller than the minimum of $\sigma(f'_j)/f_j$.

However, there are physiological limits to the precision of a temporal representation of frequency (Rose et al 1967, 1971), so that the relationship between $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ predicted above will be true only for signal levels below the levels causing saturation of the fibres, or where an external noise whose power is greater than the equivalent power of the internal noise is present. For signal levels and external signal-to-noise ratios above these limits, when both the channels tuned to the individual components and those tuned to intermediate frequencies are saturated, the signal-to-noise ratios in all these channels will be similar, so that if the spread of neural activity allows an independent representation of the component frequencies, $\sigma(f'_0)/f_0$ may be expected to be approximately equal to $\sigma(f'_j)/f_j$. If however, the spread of activity is such that the channels carrying frequency information for the individual components are also affected by other components, $\sigma(f'_j)/f_j$ will be increased, while if f'_0 is based upon residual periodicity information, such interaction is likely to lead to a decrease in $\sigma(f'_0)/f_0$. In this instance, it must be expected that $\sigma(f'_0)/f_0$ would be smaller than $\sigma(f'_j)/f_j$. This argument would also fail if the auditory filters were found to be either grossly asymmetric, or to have a constant attenuation over much of their pass-band. It has been

suggested that the auditory filter is somewhat asymmetric (Patterson and Nimmo-Smith, 1980), but not to such a degree as to invalidate the above argument.

If the auditory filters were Gaussian in form, and the bandwidth of these filters were known, the effective signal-to-noise ratio for the representation of periodicity could be estimated. Consider a filter whose centre frequency lies at 900 Hz, with an equivalent rectangular bandwidth of 160 Hz, which is Scharf's (1970) estimate of the critical bandwidth at 1000 Hz, and a signal whose components are separated by 200 Hz about a mean of 900 Hz. The attenuation of each of the powers of the two components relative to a component at 900 Hz will then be given by,

$$y(100/(160/\sqrt{2\pi})) / y(0), \quad (2.9)$$

where $y(z)$ is the ordinate of the Gaussian distribution, and the value $100/(160/\sqrt{2\pi})$ is the abscissa of the component at 800 Hz in units of the standard deviation of the distribution. The power attenuation factor is then approximately 0.298, or -5.3 dB, and taking the power of the two components together, the attenuation relative to a single component at 900 Hz will be -2.3 dB. Hence, $\sigma(f'_0)/f_0$ for a given signal-to-noise ratio would be expected to be approximately equal to $\sigma(f'_j)/f_j$ for a signal-to-noise ratio 2.3 dB smaller. A smaller bandwidth will result in a larger attenuation factor; for a rectangular bandwidth of 90 Hz, which is Houtgast's (1974) estimate of the critical bandwidth at 1000 Hz from non-simultaneous masking studies, the attenuation factor for equivalence of $\sigma(f'_0)/f_0$ and $\sigma(f'_j)/f_j$ is 14 dB. Clearly,

the quantitative relationship between these two standard deviations is critically dependent upon the estimated filter width, and it also depends upon the assumed form of the filter. Since the available estimates of the bandwidth of the auditory filter are rather variable, there is little point in attempting quantitative estimates of the relationship between these two parameters expected from a periodicity theory.

Licklider's (1959) autocorrelation theory proposes the use of the same stimulus information as does Schouten's periodicity theory, and the predictions of Licklider's theory will, if only unanalysed residual periodicity is utilized, be very similar to those of Schouten's theory.

Estimates of the Residue Pitch Difference Limen.

The frequency or pitch difference limen (DL), generally defined as the frequency or pitch difference which can be detected with a probability of 0.75 in a Two-Alternative-Forced-Choice (2AFC) task, can, as is shown in Chapter 3, be considered as an estimate of the standard deviation of an internal representation of frequency or pitch. A few published studies have investigated the difference limen for residue pitch. Ritsma (1963) employed the Method of Adjustment to estimate the DL of a harmonic SAM signal, and reported that the relative DL (the DL as a proportion of the fundamental frequency) of this signal was larger by a factor of about 5 than that for a pure tone in the region of the complex signal's components. The results of Ritsma's study are shown in

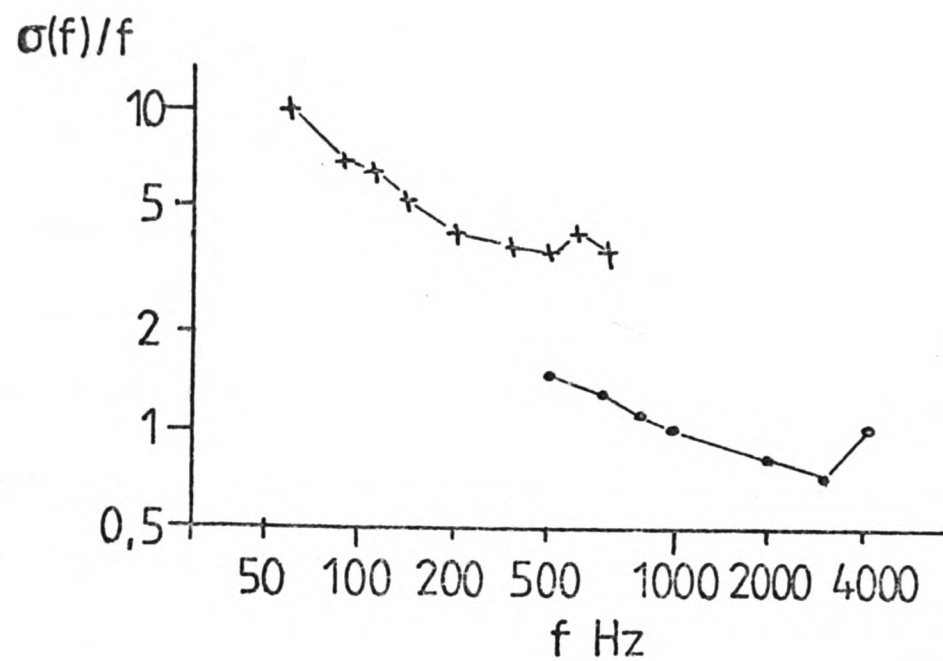


Figure 2.6 Ritsma's (1963) estimate of the DL for the residue pitch of a harmonic SAM signal as a function of fundamental frequency. The data are the standard deviations of the fundamental frequency of a second SAM signal which was adjusted by the observer to match the pitch of the first. The harmonic numbers of the carrier frequency for the two signals were 7 and 10. Also shown is the relative frequency DL for a sinusoid estimated similarly.
(+ residue pitch, • sinusoid frequency).

figure 2.6; this estimate of $\sigma(f'_0)/f_0$ varies from about 0.004 to 0.010, which, in terms of the spectral-pattern recognition models corresponds to a value for $\sigma(f'_j)/f_j$ of between about 0.002 and 0.006. The pure tone relative DL in the region of the signal components was between 0.0007 and 0.0015, which is too small to account for the relative DL of the residue pitch. Henning and Grosberg (1968) measured the DL for a pulse train, and found that for fundamental frequencies below 2000 Hz, this relative DL was smaller than that for a sinusoid at the fundamental frequency and comparable to the relative DL for single sinusoids between 1000 and 2000 Hz. A similar study was reported by Walliser (1969b) who reported results which are comparable to Henning and Grosberg's, and are illustrated in figure 2.7. The findings of Henning and Grosberg and Walliser are more consistent with the spectral-pattern recognition theories than are Ritsma's, but the relative DL for a single sinusoid may not constitute a good estimate of $\sigma(f'_j)/f_j$, since there may well be mutual masking between the components of a complex signal, which would increase $\sigma(f'_j)/f_j$ to a value greater than that of the unmasked relative DL at f_j .

Goldstein et al (1978) have estimated $\sigma(f'_j)/f_j$ for f_j between 1 and 8 KHz by fitting Goldstein's model to performance in a task requiring the identification of the residue pitch of two-component harmonic signals. Figure 2.8 shows these estimates as a function of frequency together with pure tone relative DLs from Moore (1973a) and Weir, Jesteadt and Green (1977). The estimates of $\sigma(f'_j)/f_j$ are about four times larger than the pure tone relative DLs, but the frequency

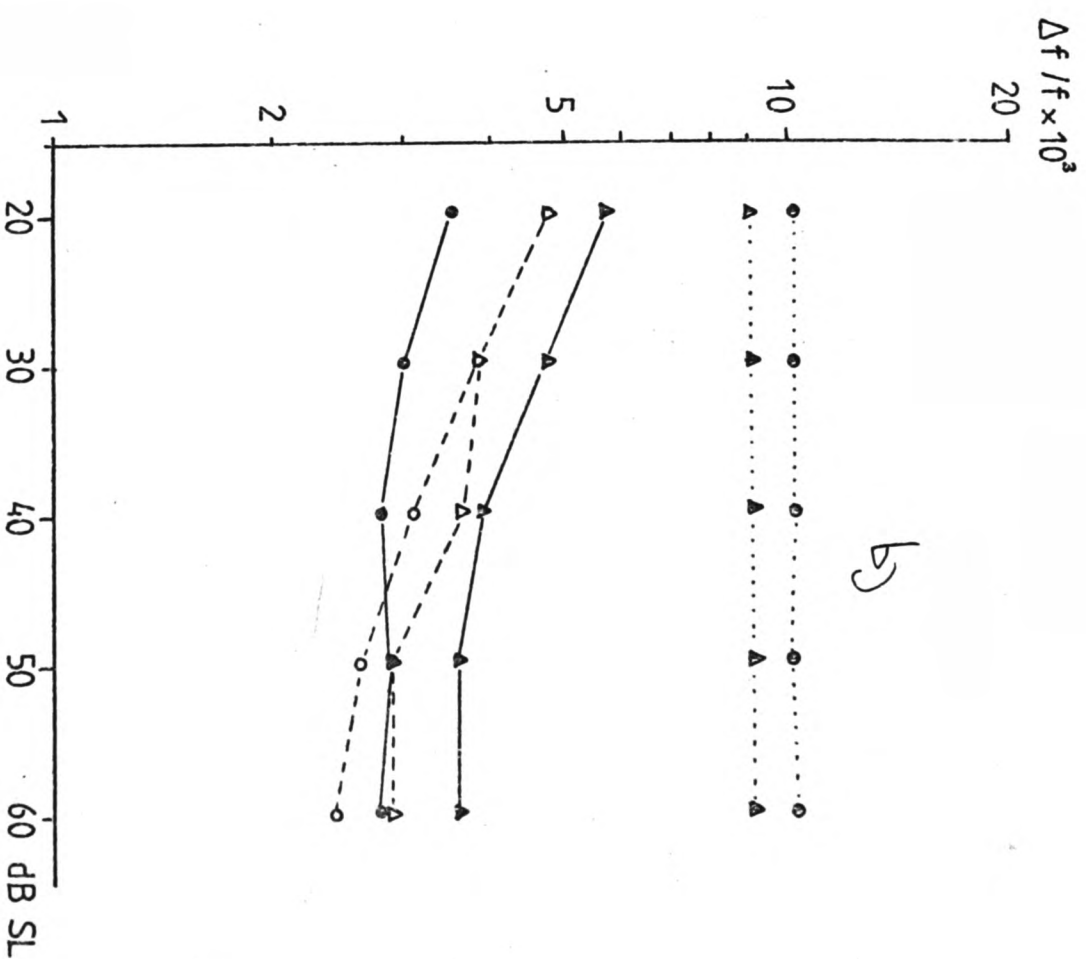
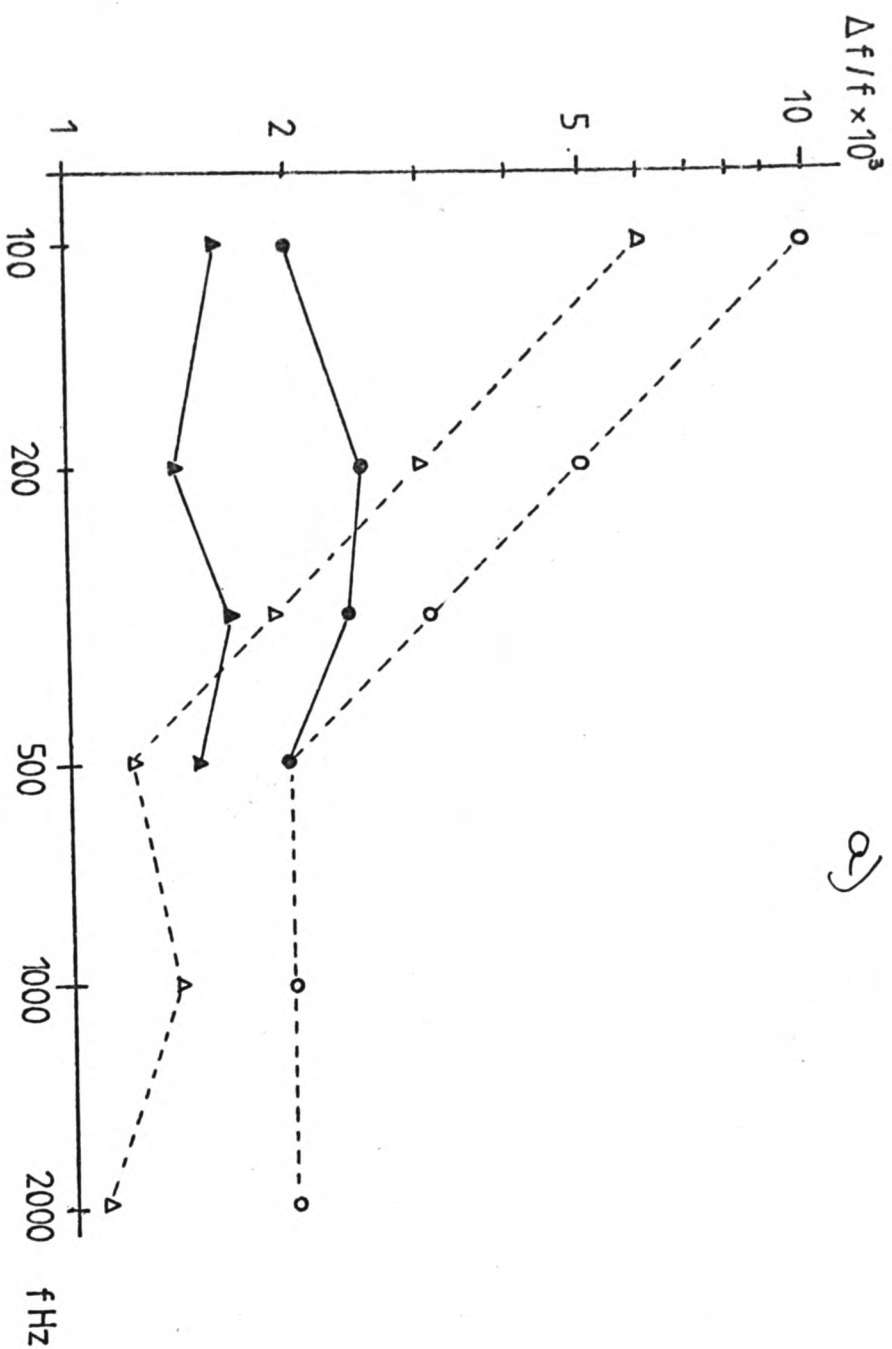
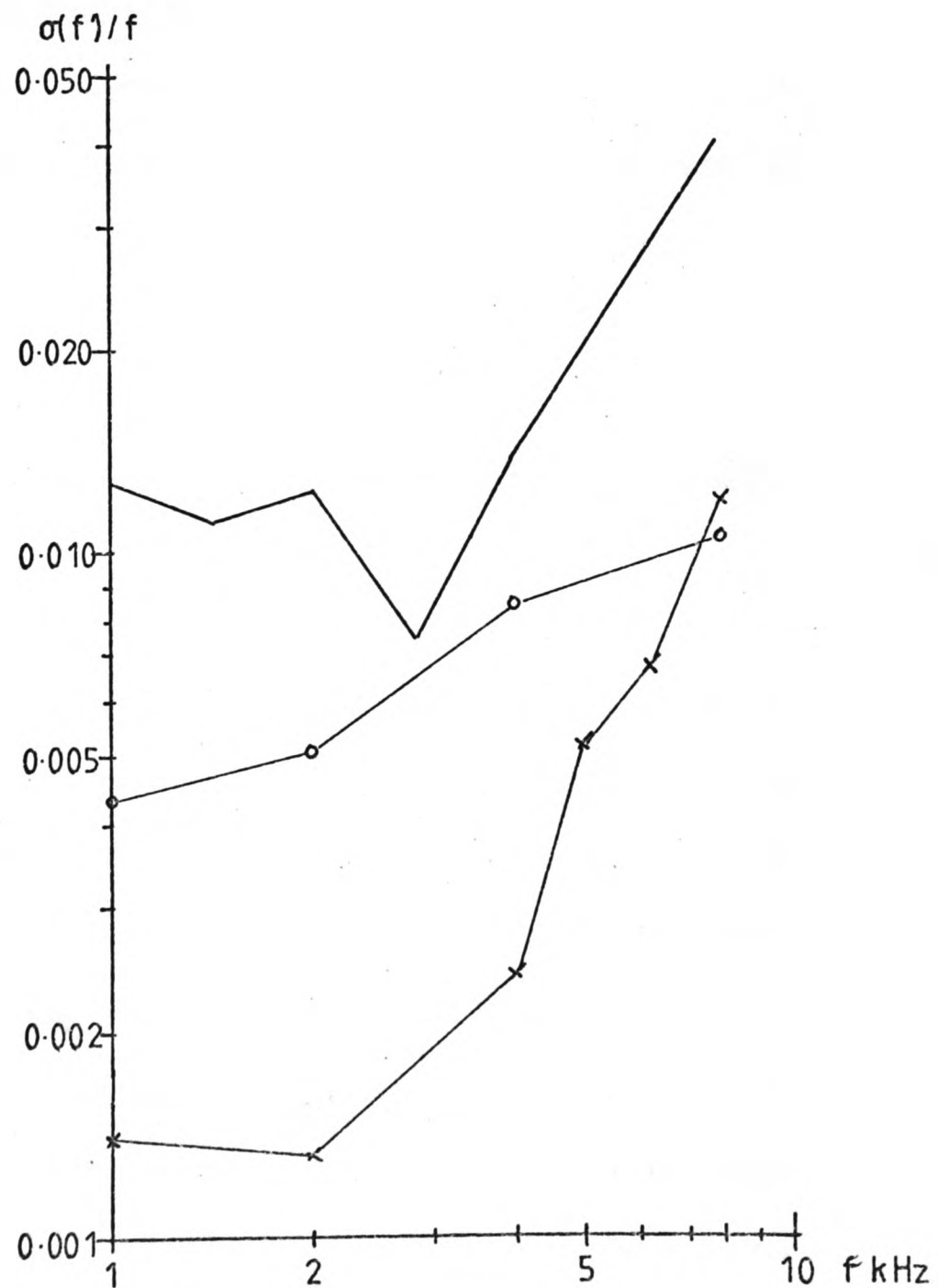


Figure 2.7 Pitch discrimination for a pulse train and a sinusoid as a function of frequency - panel a) - and sensation level - panel b) - results from Walliser (1969) for two observers. The empty symbols refer to sinusoids, and the filled symbols to pulse trains. The fundamental frequency of the pulse train for the data in panel a) was 100 Hz, and the sinusoids here were at 100 (dotted line) and 1000 Hz (solid line). The triangle and circle symbols represent the two observers.

Figure 2.8. Estimates of $\sigma(f')/f'$ as a function of frequency, from Goldstein et al (1978) - shown as plain line. Also shown are estimates of the relative frequency difference limen as a function of frequency. The open circles are data from Weir, Jesteadt, and Green (1977) for a 100 msec signal at 20 dB SL. The crosses are data from Moore (1973a) for signals of 200 msec duration and about 50 dB SL. The signals used by Goldstein et al were 500 msec in duration, and at about 20 dB SL.



dependence of the estimates of $\sigma(f'_j)/f_j$ is comparable to that for the pure tone measures, which suggests that $\sigma(f'_j)/f_j$ is at least partially determined by the same limits as those which determine the pure tone DLs.

The experiments reported in Chapters 5 to 8 below are intended to estimate, for a two-component signal, both $\sigma(f'_j)/f_j$, which is equivalent to the relative difference limen for residue pitch, and also $\sigma(f'_0)/f_0$, which is equivalent to the difference limen for a single component of a complex tone presented together with a second component. A comparison of these estimates will permit a quantitative test of the spectral-pattern recognition theories against the periodicity theories.

Computer Simulation methods for Goldstein's and Wightman's models.

The simulation procedures adopted for the derivation of the predictions of these two spectral recognition models result in an estimate of the probability density function for the pitch of the signal. Houtsma (1979) has described a simulation procedure for Goldstein's model which is similar to that employed here, and also a simulation method for Wightman's autocorrelation model. Houtsma's approach to the simulation of the latter model differs considerably from that adopted here, in that he assumes that the probability of the selection of a pitch corresponding to a peak in the autocorrelation function is directly proportional to the height of that peak. Houtsma gives no

justification for this assumption, which leads to the prediction of a considerable number of pitch modes, each having similar probabilities of selection. Since the relative heights of the autocorrelation peaks are approximately independent of the variability of component frequency representation, the probabilities of the corresponding pitches according to this assumption would also be approximately independent of this variability. The simulation method described below does take account of the dependence of the variability of the autocorrelation function on the variability of component frequency estimates, and leads to rather different predictions. In particular, the method adopted here predicts very little pitch ambiguity with a small $\sigma(f'_j)/f_j$ (<0.008), and it appears that Houtsma's method is erroneous. For this reason, Houtsma's (1979) claim that Wightman's model cannot account for the distribution of pitch values for a two-component complex, while Goldstein's model can do so, must be rejected.

In the procedures adopted here, the signal is represented stochastically. The spectrum of the signal is taken to be a sum of k Gaussian distributions, which represent internal estimates of the spectrum of each resolved component

$$|F(f)|^2 = \sum_{j=1}^k f'_j, \quad (2.10)$$

where the signal is of k components, each represented by the Gaussian population distribution f'_j , with mean $\mu(f'_j)$ and standard deviation $\sigma(f'_j)$.

The simulation procedures for Goldstein's and Wightman's models are shown as flow charts in figures 2.9 and 2.10

Figure 2.9. Simulation of Goldstein's (1973) Optimum-Processor Model.

Given data:

- Population distribution of signal amplitude spectrum $F|f|$, described by the sum of k normal distributions with means $\mu(f'_j)$ and standard deviations $\sigma(f'_j)$
- Upper and lower limits for the estimated lowest harmonic number n ,
- N , the number of trials to be simulated,
- Random number generator giving samples from a rectangular distribution of random numbers ϵ_i , where $0 < \epsilon < 1$.

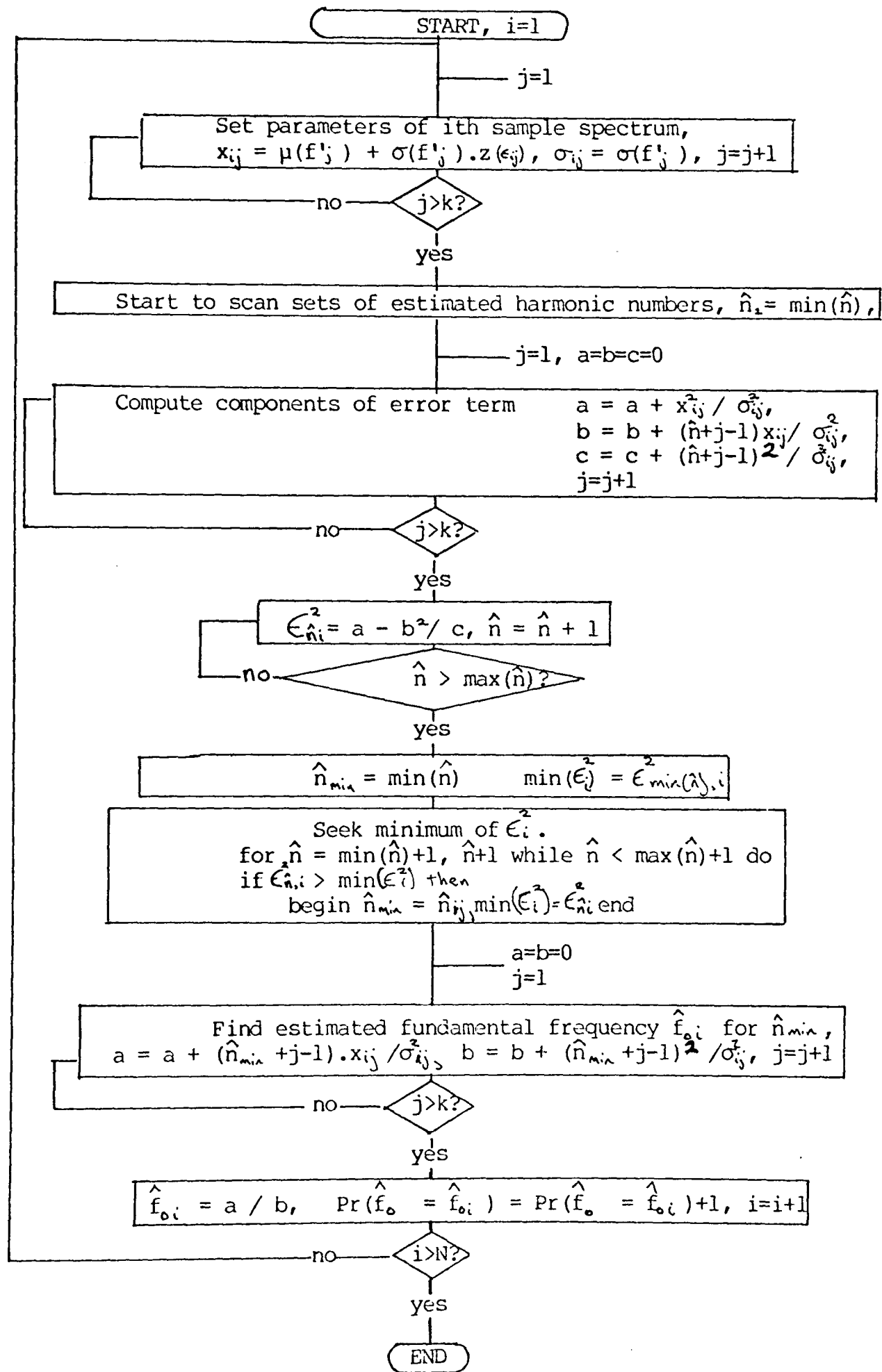
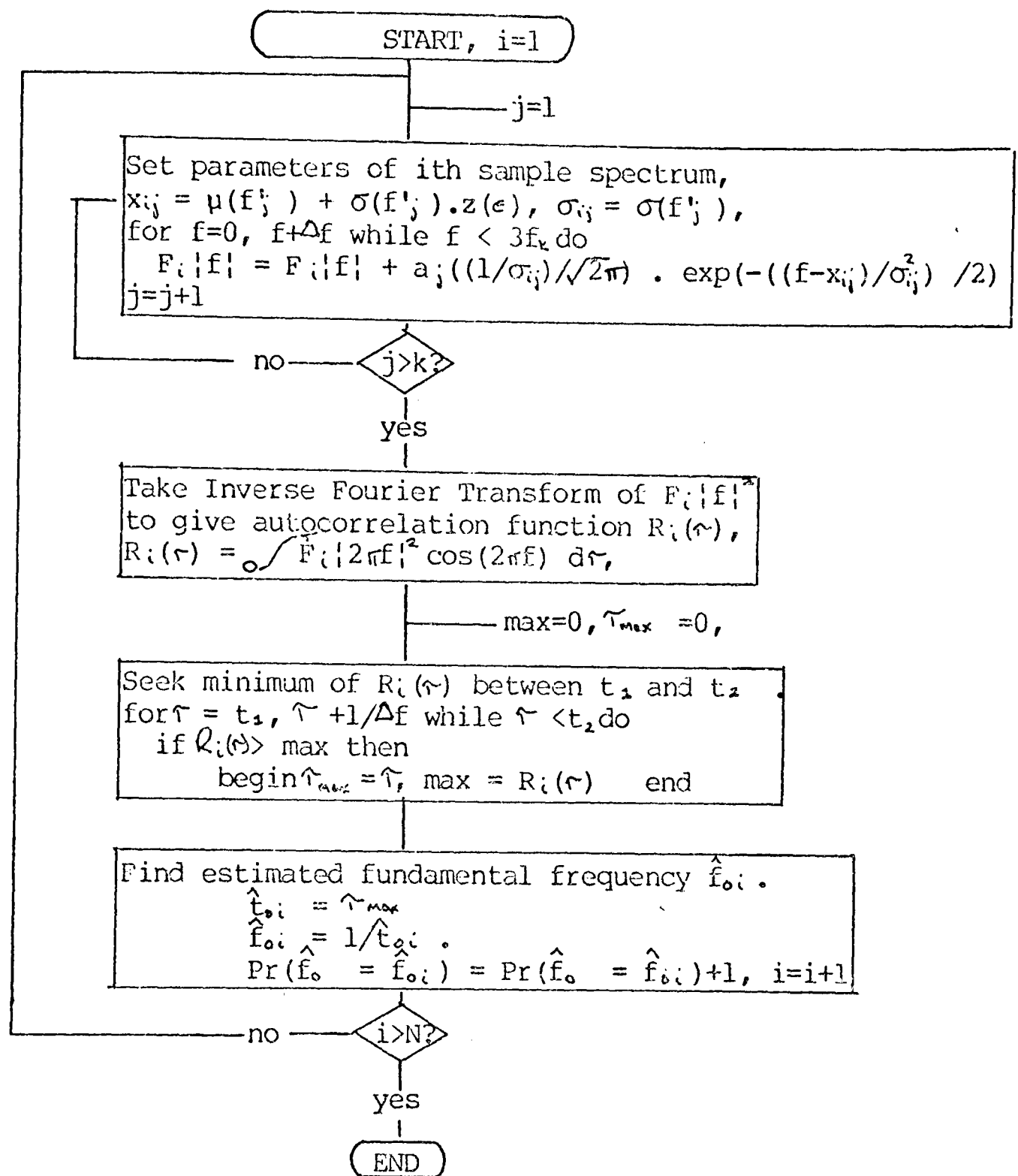


Figure 2.10. Simulation of Wightman's (1973b) Autocorrelation Model.

Given data:

- Population distribution of signal power spectrum $F|f|$, described by the sum of k normal distributions with means $\mu(f'_j)$ and standard deviations $\sigma(f'_j)$
- Upper and lower limits for the estimated period \hat{t}_0 ,
- N , the number of trials to be simulated,
- Δf , the frequency difference between successive samples of the spectrum.
- Random number generator giving samples from a rectangular distribution of random numbers ϵ , where $0 < \epsilon < 1$.



respectively. A large number of samples of each signal were taken ($N = 400$). The probability distribution for the pitch of the signal was estimated by cumulating the distribution of the best estimates of \underline{f}'_0 on each trial. Thus, on the i th trial, the sample power spectrum is the i th sample from the population distribution of the power spectrum,

$$F_i |f|^2 = \sum_{j=1}^K x'_{ij} \quad , \quad (2.11)$$

where x'_{ij} is a sample distribution taken from the population distribution \underline{f}'_j , having the mean \underline{x}'_{ij} and standard deviation $\sigma(\underline{f}'_j)$. The simulation of Wightman's model made use of an Fast-Fourier-Transform procedure (Note 2) based upon 512 or 1024 points. These simulations were run on an ICL 2980 computer and written in a dialect of Algol 60 (Note 3). Goldstein's model was simulated using a DEC PDP/8a computer; the simulation programmes were written in a second dialect of Algol 60 (Note 4).

The majority of the experiments described below employ harmonic signals having two components at the n th and $n+1$ th harmonics of 200 Hz, where $n = 4$ or 5. The simulation procedures for Goldstein's and Wightman's models described above have been applied to these two signals, and in figures 2.11 and 2.12 the predicted areas of the various residue pitch modes for these two signals are shown as a function of $\sigma(\underline{f}'_j)/f_j$. The predictions of the two models differ only slightly, the only major difference being that Wightman's model gives a greater predicted likelihood of higher pitches where $\sigma(\underline{f}'_j)/f_j$ is larger than about 0.020. The experiments described in Chapter 10 are intended to estimate the distribution of pitch matches to

Figure 2.11 The areas of the theoretical pitch modes for Goldstein's model as a function of $\sigma(f_j)/f_j$. These results are for a two-component harmonic signal, containing the 4th and 5th harmonics - panel a) - and the 5th and 6th harmonics - panel b) - of 200 Hz. The results were obtained by the computer simulation method described in the text.

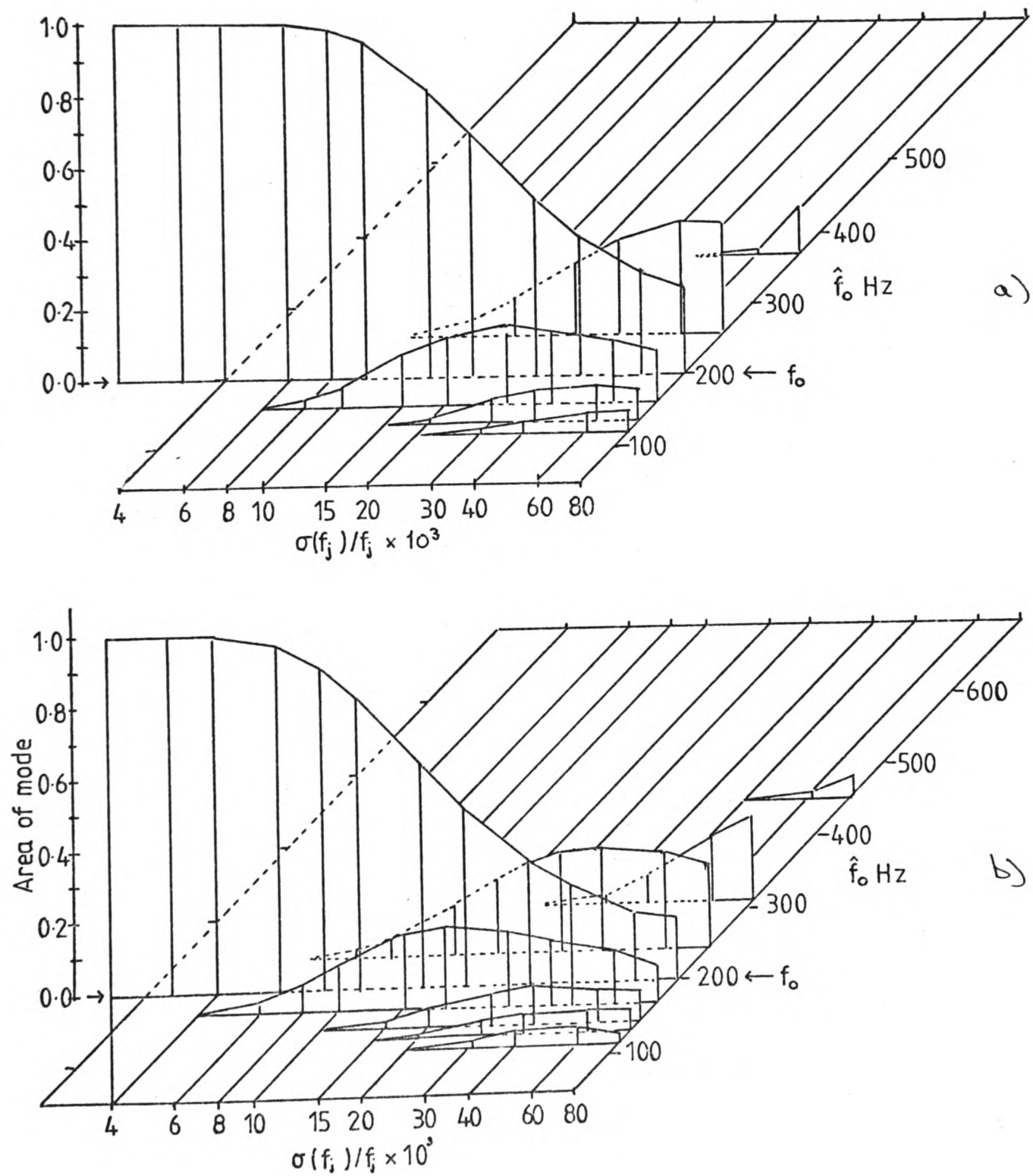
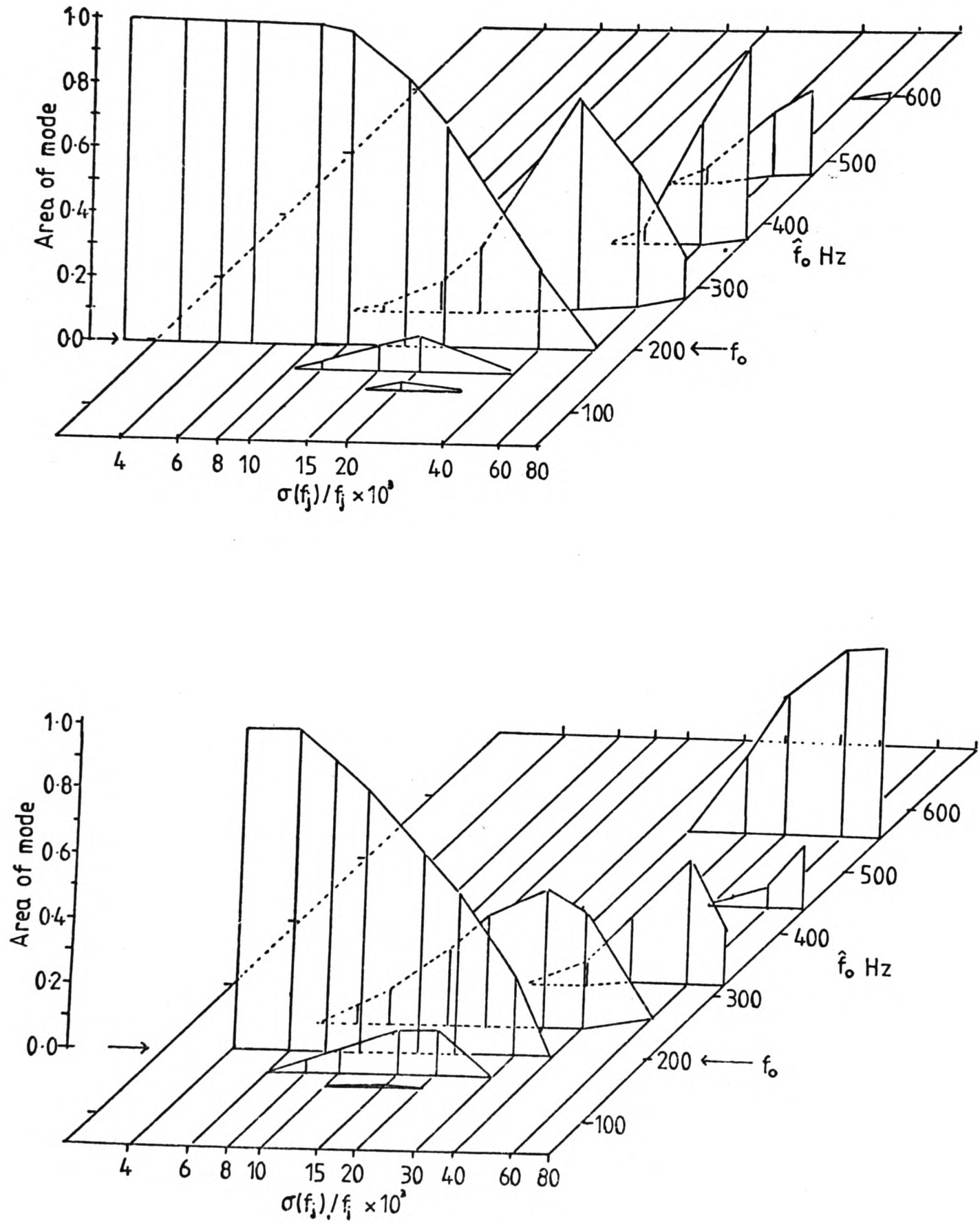


Figure 2.12 The areas of the theoretical pitch modes for Wightman's model as a function of $\sigma(f_j)/f_j$. Panel a) shows simulated results for a signal containing the 4th and 5th harmonics of 200 Hz, and panel b) the results for a signal containing the 5th and 6th harmonics of 200 Hz.



these signals by human observers as a function of the presentation conditions of the signal. Since psychophysical estimates of $\sigma(f'_j)/f_j$ for the same signals and presentation conditions are available from the experiments described in Chapters 5-8, these pitch matches may be compared with the theoretical distributions derived from the computer simulations of Wightman's and Goldstein's models. Houtsma and Goldstein (1972) and Houtsma (1979) have investigated musical interval recognition performance for the residue pitch of two-component harmonic signals, and have reported their results in terms of the probability of assigning a residue pitch at the fundamental frequency. From these data, obtained for signals with a range of harmonic numbers, they have fitted to Goldstein's model a theoretical value of $\sigma(f'_j)/f_j$ as a function of f_j . They have not, however, attempted a direct estimation of $\sigma(f'_j)$ or $\sigma(f'_o)$ from a pitch discrimination task.

Since the relation between $\sigma(f'_j)/f_j$ and $\sigma(f'_o)/f_o$ for Goldstein's model is known, this relation can be used as a check for the correct operation of the simulation procedures applied to that model. Figure 2.13 presents scatter-plots of the simulated values of $\sigma(f'_o)/f_o$ as a function of $\sigma(f'_j)/f_j$, together with the values predicted from equation (2.7). In the case of Wightman's model, this relationship has not been derived, and is thus estimated from computer simulations. Figure 2.14 presents a scatter-plot of the relation between $\sigma(f'_j)/f_j$ and the estimate of $\sigma(f'_o)/f_o$ derived from these simulations. For two-component signals where both

Figure 2.13 Simulated values of $\sigma(f'_0)/f_0$ as a function of $\sigma(f'_j)/f_j$ for Goldstein's model. The solid line is the theoretical relationship. The simulation results are close to the theoretical values except for $\sigma(f'_j)/f_j$ less than 0.005, where the resolution of the simulation procedure dominates.

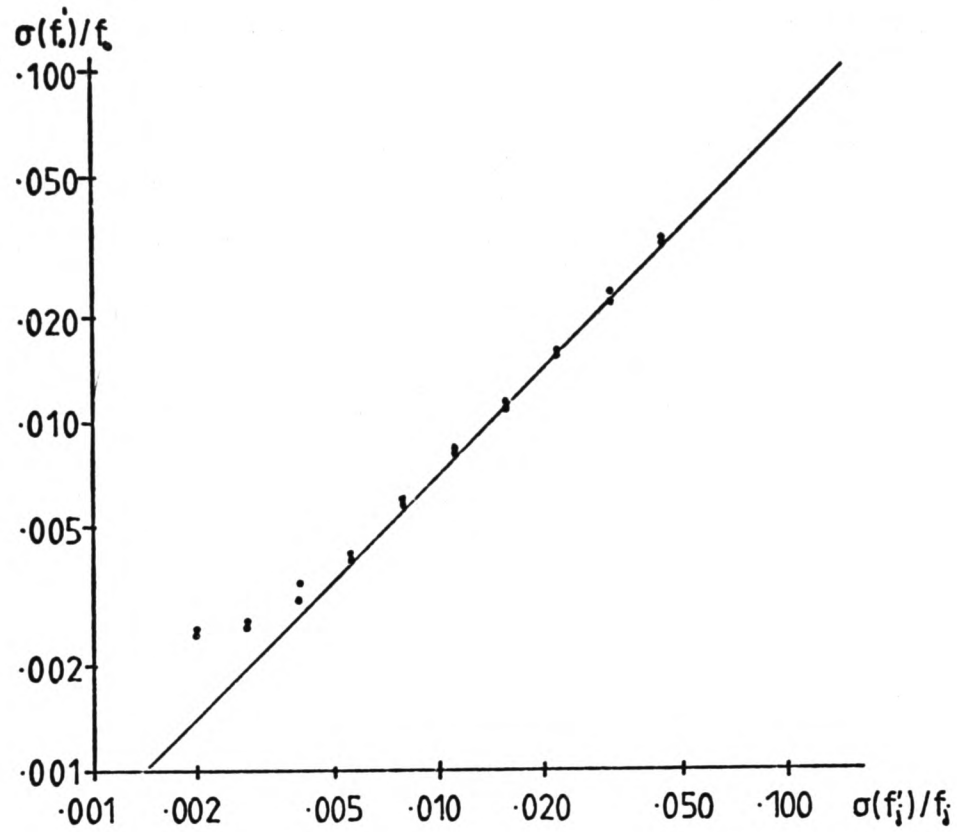
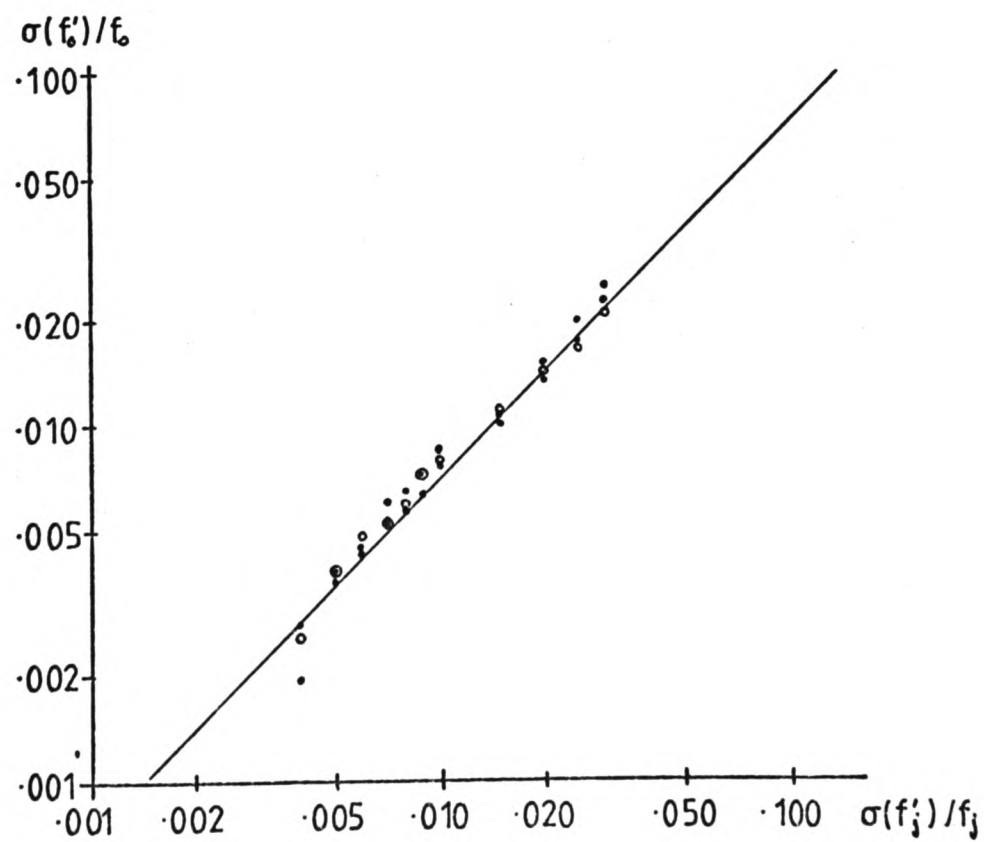


Figure 2.14 Simulated values of $\sigma(f'_0)/f_0$ as a function of $\sigma(f'_j)/f_j$ for Wightman's model. The solid line is the relationship expected from the assumption that the distribution of f'_0 around f_0 has a variance given by equation 2.7.



$\sigma(f'_j)/f_j$ are equal, the simulations result in pdfs where $\sigma(f'_0)/f_0 = \sigma(f'_j)/f_j \cdot 1/\sqrt{2}$, which indicates that the relation between these two parameters is the same for Wightman's model as for Goldstein's. The simulation results illustrated in figures 2.11 to 2.14 are appropriate to the two-component harmonic signals used for the majority of the experiments described below, where these signals consist of the 4th and 5th or the 5th and 6th harmonics of 200 Hz. Chapter 10 describes several experiments intended to provide empirical estimates of the pitch pdf for these signals.

Summary.

The spectral-pattern recognition models due to Goldstein (1973) and Wightman (1973b) have been presented formally, together with computer simulation methods by which these models can predict a probability density function for the residue pitch of an arbitrary signal. Density functions have been predicted for two harmonic signals having two components, and these may be compared with the psychophysical estimates of the pitch distribution of these signals derived as described in Chapter 10.

Where the components of a harmonic complex signal are well-resolved, a general prediction from the spectral-pattern recognition theories is that the relative pitch difference limen for the residue pitch at the fundamental frequency of a harmonic signal should be smaller than the smallest of the relative difference limens for the signal components presented together

with the other components. The reverse relationship may be predicted from a periodicity theory. The experiments described in Chapters 5 to 8 address this question.

CHAPTER 3

Abstract.

This chapter introduces the statistical basis for the estimation of the parameters $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ from a pitch discrimination task. This estimation method is based on a Signal Detection model (Green and Swets, 1966; Luce and Galanter, 1963) of a decision process for a task in which an observer is required to judge, with respect to some internal dimension, the relative values of two stimuli presented in two successive temporal intervals. The application of the model to pitch and frequency discrimination depends primarily upon the assumption that the internal representation of the frequency of a component of a complex auditory signal is a normally distributed random variable. This chapter also refers to some of the theories of frequency discrimination for simple sinusoidal signals, and considers the physical factors likely to influence the parameter $\sigma(f'_j)/f_j$. The theories considered are an Energy Detector model (Green and Swets, *ibid*) based upon spectral analysis, which has been adapted for frequency discrimination by Henning (1967a), a temporally based model due to Siebert (1970), and modifications of that model described by Luce and Green (1974) and by Goldstein and Srulovicz (1977).

These theories are generalized to the discrimination of a residue pitch derived from several well-resolved components of a complex signal.

Introductory comments.

In the previous chapter, contrasting predictions were made from the spectral-pattern recognition models and Schouten's periodicity theory concerning the relations between the theoretical parameters $\sigma(f'_j)/f_j$, the relative standard deviations of internal estimates of the frequencies of the components of a harmonic complex signal, and $\sigma(f'_0)/f_0$, the relative standard deviation of the internal estimate of the residue pitch at the fundamental frequency of the complex. The experiments described in Chapters 4 to 8 employ the decision model described below to estimate these two parameters from psychometric functions relating discrimination performance to the relative frequency difference between the component and fundamental frequencies of two-component complexes. This model has been tested a number of times with the results from simple discrimination experiments, and has been shown to be consistent with these data (Green and Swets, 1966; Luce and Galanter, 1963), and also to be specifically applicable to frequency discrimination for single pure tones (Creelman and Macmillan, 1979).

The Signal Detection Model of Discrimination.

Let the internal representation of the frequency f_j of a single component of a complex signal be assumed to be a Gaussian random variable f'_j having the mean $\mu(f'_j)$ which is linearly dependent on f_j , and the standard deviation $\sigma(f'_j)$. If f'_j is independent of the internal representation of other components,

then $\sigma(f'_j)$ then may be directly estimated from the psychometric function relating performance in a Two-Alternative-Forced-Choice (2AFC) discrimination task to the frequency difference between two signals at the frequencies f_j and $f_j + \Delta f_j$. In such a task, an observer attempts to order the frequencies of two pulsed sinusoids at the frequencies f_j and $f_j + \Delta f_j$, where Δf_j is small with respect to f_j .

The Theoretical Psychometric Function.

The psychometric function can be illustrated by reference to an idealized 2AFC experiment.

Two stimuli x_1 and x_2 differing with respect to a single physical dimension x , which are presented during two successive temporal intervals, and have the normally distributed internal representations x'_1 and x'_2 , may be considered to give rise, on a given trial, to the sample estimates \hat{x}_1 and \hat{x}_2 drawn from the two population distributions \hat{x}'_1 and \hat{x}'_2 . The probability that an observer will judge the stimulus x_2 to be greater than x_1 with respect to x on a given trial may be assumed to depend upon the difference between the two sample estimates. Since both x'_1 and x'_2 are, by assumption, normally distributed, the distribution of $(x'_2 - x'_1)$ is also normal, with the mean

$$\mu(x'_2 - x'_1) = \mu(x'_2) - \mu(x'_1) \quad (3.1)$$

and the variance

$$\sigma^2(x'_2 - x'_1) = \sigma^2(x'_1) + \sigma^2(x'_2) - 2 r(x'_1, x'_2) \cdot \sigma(x'_1) \cdot \sigma(x'_2), \quad (3.2)$$

where $r(x'_1, x'_2)$ is the correlation of the distributions

\underline{x}'_1 and \underline{x}'_2 . The theoretical internal distributions \underline{x}'_1 and \underline{x}'_2 , and the distribution of their differences $\underline{x}'_2 - \underline{x}'_1$, are illustrated in figure 3.1. The probability that the sample difference $\hat{\underline{x}}_2 - \hat{\underline{x}}_1 > 0$ is given by

$$p(\hat{\underline{x}}_2 - \hat{\underline{x}}_1 > 0) = 1 - \Phi[\mu(\underline{x}'_2 - \underline{x}'_1)/\sigma(\underline{x}'_2 - \underline{x}'_1)], \quad (3.3)$$

where

$$\Phi(y) = 1/\sqrt{2\pi} \int_{-\infty}^y \exp[-(x^2/2)] dx, \quad (3.4)$$

the area under the normal density function from $-\infty$ to y . From equation 3.1, $\mu(\underline{x}'_2 - \underline{x}'_1)$ will be a monotonic function of the physical difference $\underline{\Delta x}$ between the stimuli \underline{x}_2 and \underline{x}_1 . If $\underline{\Delta x}$ is equally often positive and negative, the absolute value of $\underline{\Delta x}$ is constant, and $\mu(\underline{x}'_i)$ is a linear function of \underline{x} , then an unbiased observer will equally often make the responses $R(\underline{x}_2 > \underline{x}_1)$ and $R(\underline{x}_2 < \underline{x}_1)$, where $R(\underline{x}_i > \underline{x}_j)$ indicates the judgement that stimulus \underline{x}_i is greater than stimulus \underline{x}_j with respect to \underline{x} , and the response $R(\underline{x}_i < \underline{x}_j)$ indicates the reverse judgement. The response of such an observer may be said to be based upon a symmetric decision rule, specified by

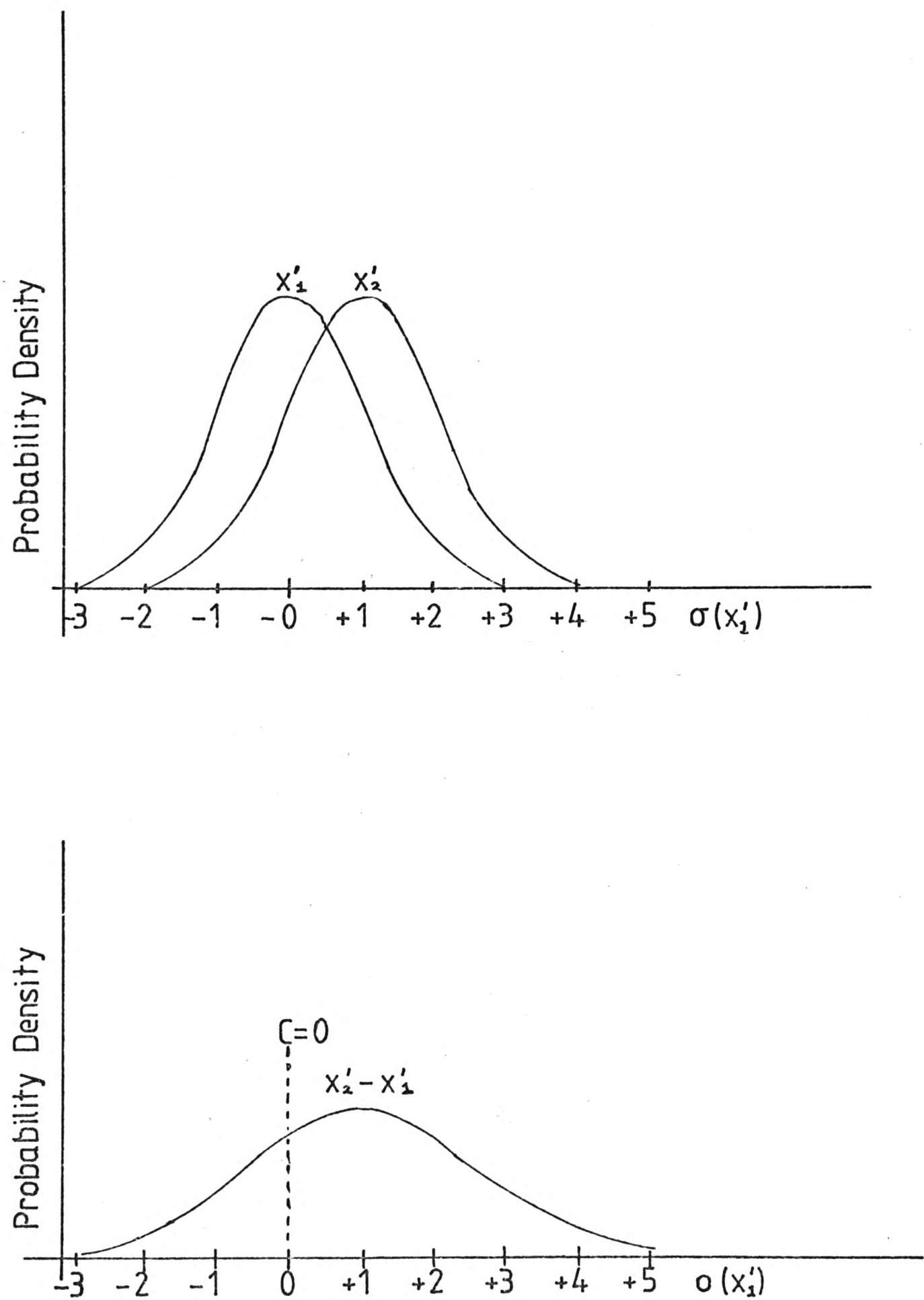
$$\begin{aligned} R &= R(\underline{x}_2 > \underline{x}_1) \text{ if } \hat{\underline{x}}_2 - \hat{\underline{x}}_1 > 0, \text{ and} \\ R &= R(\underline{x}_2 < \underline{x}_1) \text{ if } \hat{\underline{x}}_2 - \hat{\underline{x}}_1 < 0. \end{aligned} \quad (3.5)$$

The probability that an unbiased observer will make the response $R(\underline{x}_2 > \underline{x}_1)$ is thus the probability that $\hat{\underline{x}}_2 - \hat{\underline{x}}_1 > 0$. Since the difference $(\underline{x}'_2 - \underline{x}'_1)$ is a continuously distributed random variable, the probability that $\hat{\underline{x}}_2 - \hat{\underline{x}}_1 = 0$ is zero.

A biased observer will respond according to an asymmetric decision rule, such that

$$\begin{aligned} R &= R(\underline{x}_2 > \underline{x}_1) \text{ if } \hat{\underline{x}}_2 - \hat{\underline{x}}_1 > C, \text{ and} \\ R &= R(\underline{x}_2 < \underline{x}_1) \text{ if } \hat{\underline{x}}_2 - \hat{\underline{x}}_1 < C, \end{aligned} \quad (3.6)$$

Figure 3.1 The postulated distributions of the two stimuli in the 2AFC task and the distribution of their difference. All the distributions are scaled in units of $\sigma(x_1)$.



where C represents a criterion with some non-zero value: The probability that a biased observer will make the response $R(x_2 > x_1)$ will thus be the probability that $\hat{x}_2 - \hat{x}_1 > C$.

The psychometric function is the function relating the probability of some response, such as $R(x_2 > x_1)$, to Δx or some monotonic function $g(\Delta x)$ of Δx , and will be the normal ogive corresponding to the integral of the distribution of $(x'_2 - x'_1 - C)$, whose mean is equal to $\mu(x'_2 - x'_1) - C$, and whose variance is as defined by equation 3.2. The constant term C will be omitted from the remainder of this discussion.

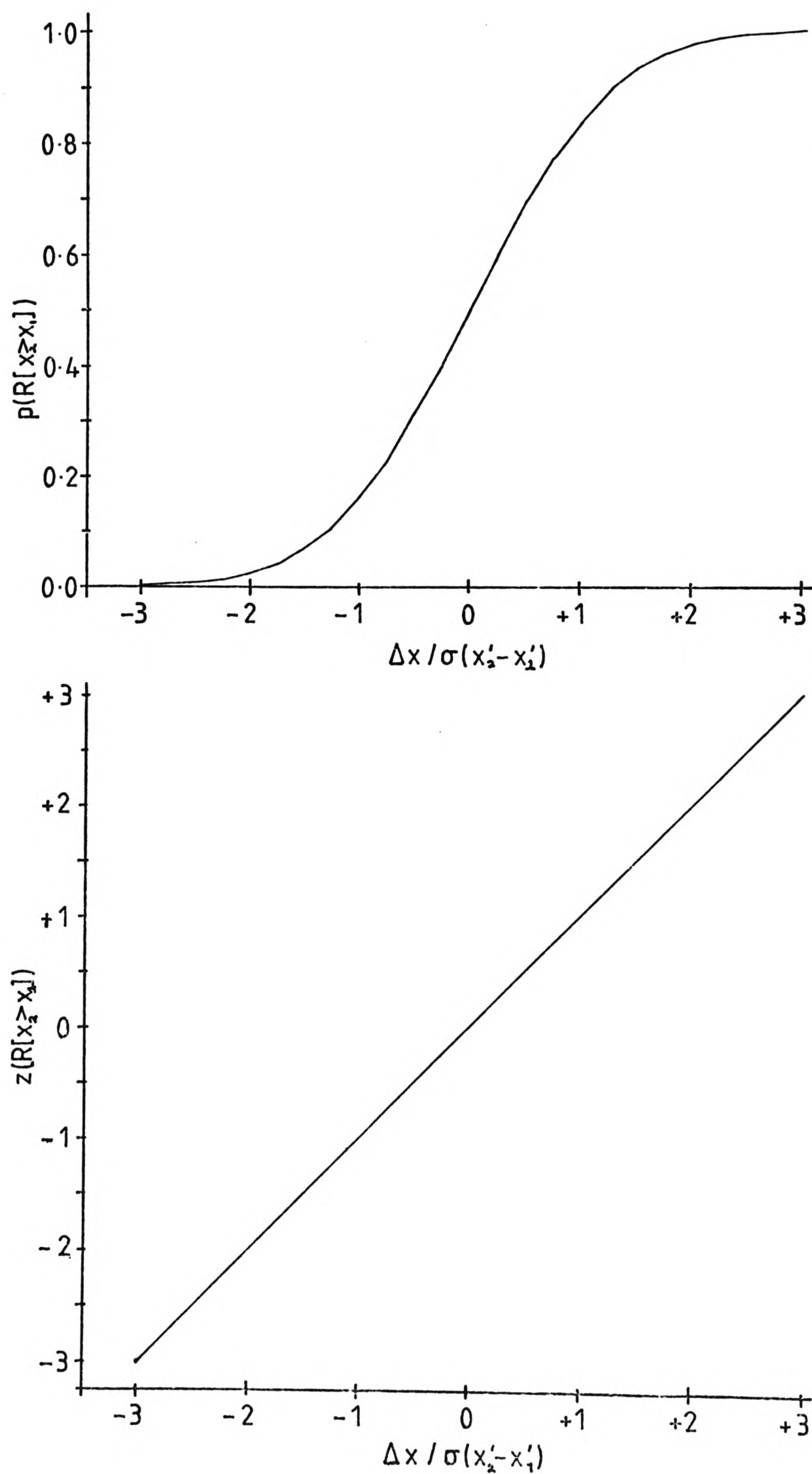
If it can be reasonably assumed that the distributions x'_1 and x'_2 are independent, that is, there is no common component in $\sigma(x'_1)$ and $\sigma(x'_2)$, then $r(x'_1, x'_2) = 0$, so that from equation 3.2,

$$\sigma^2(x'_2 - x'_1) = \sigma^2(x'_1) + \sigma^2(x'_2). \quad (3.7)$$

If it is also assumed that either the variances $\sigma^2(x'_1)$ and $\sigma^2(x'_2)$ are equal, or that one of these variances is known, the psychometric function contains all the information required for the estimation of the two equal variances or the single unknown variance.

When the theoretical probability of a response such as $R(x_2 > x_1)$ is standardized according the normal 'z' deviates, the theoretical psychometric function will be a linear function of Δx or $g(\Delta x)$, which will have a intercept at $\Delta x=0$ if the observer is unbiased, and a slope of $1/\sigma(x'_2 - x'_1)$. Figure 3.2 illustrates a psychometric function with both probability and

Figure 3.2 The theoretical psychometric function for an unbiased observer. The upper panel shows a theoretical function for response probability as a function of Δx , while the lower panel shows the linear function for the standardized z transform of probability.



standardized ordinates.

The Estimated Psychometric Function.

An estimated psychometric function may be obtained from an experiment which provides estimates of the probability of some response, such as $R(x_2 > x_1)$, as a function of Δx . If the estimated psychometric function can be matched to a theoretical function, which requires that either the distribution $(x'_2 - x'_1)$ is normal on a linear scale of Δx , or that $g(x)$ is known, and also that the observed function does not differ significantly from the form of the theoretical function, then the slope and intercept of the standardized observed function can be estimated. Hence, $\sigma(x'_2 - x'_1)$ is estimated as the reciprocal of the estimated slope of the observed function. If $\sigma(x'_1) = \sigma(x'_2)$, then $\sigma(x'_1)$ may be found from equation 3.2 to be

$$\sigma(x'_1) = \sqrt{\sigma^2(x'_2 - x'_1) / 2}, \quad (3.8)$$

and if $\sigma(x'_2) \neq \sigma(x'_1)$, but $\sigma(x'_1)$ is known, $\sigma(x'_2)$ may be found from equation 3.2 to be

$$\sigma(x'_2) = \sqrt{\sigma^2(x'_2 - x'_1) - \sigma^2(x'_1)}. \quad (3.9)$$

If the variances $\sigma^2(x'_1)$ and $\sigma^2(x'_2)$ are equal, and the observer is unbiased, $\sigma(x'_1)$ stated in units of x or $g(x)$ will be equal to that value of Δx at which $p(c)$, the probability of a correct response, is given by

$$p(c) = (1/\sqrt{2}) \approx 0.76, \text{ where}$$

$$p(c) = p(R[x_2 > x_1] \mid x_2 > x_1)$$

$$= p(R[x_2 < x_1] \mid x_2 < x_1). \quad (3.10)$$

In such a case, the estimate of $\sigma(x'_1)$ is thus approximately equal to the difference limen (DL), which is conventionally defined as that value of Δx at which $p(c) = 0.75$.

The Validity of the Statistical Assumptions.

The reasonableness of some of the above assumptions can be assessed from the results of such an experiment.

Normality of the Distribution ($x'_2 - x'_1$).

The form of the theoretical psychometric function describing performance as a function of Δx or $g(\Delta x)$ will, if the distribution of $x'_2 - x'_1$ is normal, be that of a cumulative normal ogive, and the identity of the forms of a theoretical and an observed function can be tested within the limits of binomial variability by the statistic χ^2 , although such identity does not strictly imply that x'_2 and x'_1 are also normal. The assumption that the observer is unbiased can be accepted if the estimated intercept of the observed function does not differ significantly from 0. This assumption, however, is not essential for the estimation of $\sigma(x'_1)$, but merely simplifies the derivation, since the intercept of the standardized observed function need not be estimated.

The Equality of the Variances $\sigma^2(x'_1)$ and $\sigma^2(x'_2)$.

If the stimuli x'_1 and x'_2 are identical in all respects

apart from the difference Δx and the time at which they are presented, it is also reasonable to assume that $\sigma^2(x'_1) = \sigma^2(x'_2)$. In other cases, this will not be reasonable, as for example, where the stimuli x_1 and x_2 differ with respect to dimensions other than x . Here, an estimation of $\sigma(x'_2)$ requires the prior estimation of $\sigma(x'_1)$ from an experiment with a pair of stimuli of the form of x_1 .

The Independence of x'_1 and x'_2 .

The assumption that x'_1 and x'_2 are independent, that is, that $r(x'_1, x'_2)$, the correlation of the distributions x'_1 and x'_2 , is zero, is difficult to test empirically, but its reasonableness depends upon the conditions of stimulus presentation. This assumption will be reasonable where the stimuli x_1 and x_2 are sinusoidal tone-bursts, or sums of such signals, and where, by the sampling theorem (Shannon and Weaver, 1949), the time interval between the onsets of the two signals is no less than $1/2f_c$, where f_c is the upper cut-off frequency of a noise process of uniform spectral density, and which is the only component of $\sigma(x'_i)$.

The application of the decision model to the estimation of $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$.

Given suitably designed experiments and the above assumptions, the parameter $\sigma(f'_j)/f_j$ can be estimated from a 2AFC frequency discrimination task. In order to estimate $\sigma(f'_j)/f_j$ for a component of a complex signal, two pulsed

sinusoids at the frequencies $f_j + \Delta f_j$ may be presented together with the other components of the complex signal, so that any effect that those other components may have upon $\sigma(f'_j)/f_j$ is included in the estimate. Two frequency discrimination experiments designed for this purpose are described in Chapters 5 and 7.

In order to extend this decision model to a 2AFC discrimination task for the residue pitch of a complex signal, two further assumptions are required. First, it must be assumed that the residue pitch derived from a complex signal has a normally distributed internal representation, an assumption which follows from the assumptions made by the spectral-pattern recognition models. The second assumption required is that this distribution is effectively unimodal. Since it is apparent that the pitch pdf for residue pitch is likely to be multimodal except where $\sigma(f'_j)/f_j$ is rather small, the estimation of $\sigma(f'_0)/f_0$ from a pitch discrimination task must be made under conditions which effectively limit the observer to sampling from a single mode of this distribution. In the experiments described below, two precautions are taken to maximize the likelihood that the observer samples from only one mode. Firstly, the fundamental frequency difference Δf_0 between the two signals compared on a trial is kept considerably smaller than the frequency difference corresponding to the pitch difference between adjacent modes of the theoretical pitch pdfs. Secondly, the mean fundamental frequency of the two signals is held constant, so that the observer always expects the residue pitch of the complex signals to be in the same small region.

Several of the experiments reported below apply this model to the estimation of $\sigma(f'_0)/f_0$, where f'_0 is the mode of the pitch pdf in the region of the true fundamental frequency f_0 of the signal.

For the experiments reported below, the estimation of the standard deviation $\sigma(x'_2 - x'_1)$ was performed by means of a Probit analysis (Finney 1971); this analysis results in a maximum-likelihood estimate of the slope and intercept of a standardized psychometric function, so that $\sigma(x'_2 - x'_1)$ can be estimated as the reciprocal of the slope of the best-fitting theoretical function resulting from this analysis.

What Factors are Likely to Affect $\sigma(f'_j)$?

While it is possible to evaluate the empirical relationship between $\sigma(f'_j)/f_j$ and the observed distribution of the observer's estimate of the residue pitch f'_0 for a single set of stimulus presentation conditions, a technique permitting the variation of $\sigma(f'_j)/f_j$ by means of the manipulation of presentation conditions represents a far more powerful method of testing the predictions of the models. Although Goldstein (1973) and Gerson and Goldstein (1978) have considered $\sigma(f'_j)/f_j$ to be limited by an internal noise and to be dependent only upon frequency, here the possibility is entertained that this parameter also depends upon those stimulus conditions which have been shown to affect frequency discrimination for simple sinusoidal signals. The results of studies of frequency discrimination are generally reported in

terms of the difference-limen (DL), that value of the frequency difference Δf which can be correctly identified as positive or negative in a 2AFC task with the probability $p(c)$ of approximately 0.75 (some studies have used criteria for the DL of $p(c) = 0.71$ or 0.76). As has been shown above, the value of Δf at which $p(c) = 0.76$ is approximately equal to the standard deviation of the internal representation of the stimulus frequency, and thus where the DL is determined from a 2AFC discrimination task, the DL is approximately equal to $\sigma(f'_j)$ where f_j is the only component of the signal. While it can not be assumed that $\sigma(f'_j)/f_j$ will be unaffected by an increase in the number of components from one to two or more, it seems unreasonable to claim that factors other than frequency which are known to affect the frequency DL for a single sinusoid will have no effect upon $\sigma(f'_j)/f_j$ for a component of a complex signal.

Frequency Discrimination for Single Sinusoids.

A number of studies of the DL for the frequency of a single sinusoid have been reported, in which a variety of parameters have been manipulated and shown to affect the DL. The effects of frequency and signal amplitude have been investigated over a wide range (e.g. Harris, 1952, Shower and Biddulph, 1931; Weir, Jesteadt and Green, 1977), as has the effect of signal-to-noise ratio at different frequencies (Henning, 1967b) and the effect of signal duration at different frequencies (Henning, 1970; Liang and Chistovich, 1961; Moore, 1973b, Turnbull, 1944). While it would be sufficient for the purposes

of this study to merely identify the parameters likely to affect $\sigma(f'_j)/f_j$ for the internal estimate of the frequency of a component of a complex signal, it is also likely to be instructive to consider attempts to describe theoretically the effects of these parameters on the DL.

Theories of auditory frequency discrimination have, as do theories of the perception of residue pitch, two parts, corresponding to a set of assumptions describing the way in which frequency is coded, and a hypothesized decision process whereby an observer's response is derived from that code. These theories further parallel each other in that there exist two general classes of theories of frequency discrimination, where frequency is assumed to be coded by either the spatial distribution of activity over a series of frequency selective channels, where such a distribution is equivalent to the noisily estimated amplitude spectrum of the peripherally processed signal, or by means of the temporal structure of the neural response to the signal in one or more channels. This dichotomy between theories of frequency discrimination can not however be simply identified with the dichotomy between the spectral-pattern recognition models and Schouten's periodicity theory, since it is perfectly possible that a spectral pattern is coded in terms of the temporal structure of the firing pattern in several neural channels each responding independently to one of the several components of a complex signal. The value of a consideration of these contrasting classes of theories of frequency discrimination is that the dependence of the acuity of frequency discrimination for the components of a complex tone

upon the physical parameters considered by the theories may be such as to allow the rejection of particular instances of these theories as inappropriate to the way in which the frequencies of the components of complex tones are encoded. The rejection of a theory as inappropriate in this case does not of course necessarily imply that the theory is also inappropriate for frequency discrimination with single sinusoids.

The Energy Detection Model.

The best known and most precisely formulated theory of frequency discrimination based upon spectral information is that invoking the energy detector (Green and Swets, 1966). The energy detector, a device well known in electrical engineering, produces an output proportional to the energy of the input signal as integrated over a period of time T_I . If the input is a function of time, $x(t)$, then the energy E of the signal is given by

$$E = \int_0^T x^2(t) dt, \quad (3.11)$$

where T is equal to T_I if the signal duration T_s is greater than T_I , and equal to the signal duration T_s if that duration is less than T_I . The application of the energy detector to frequency discrimination also requires the assumption that the input to the energy detector is the output of a band-pass filter with the transfer function $H(w)$, where w represents angular frequency. If the input signal $x(t)$ is a sinusoid, the amplitude of the signal $x'(t)$ resulting from the transformation

of $\underline{x}(t)$ by $\underline{H}(w)$ will depend upon the difference between the angular frequency \underline{w} of $\underline{x}(t)$ and \underline{w}_f , the centre frequency of the filter, and thus the detected energy \underline{E}' of $\underline{x}'(t)$ will also depend upon this frequency difference.

An Energy Detector model for frequency discrimination has been described by Henning (1967a). In this formulation, only a single filter is considered, with the Gaussian transfer function

$$H(w) = \{\exp [-(1/2) \cdot (w_f - w)^2 / F^2]\}, \quad (3.12)$$

where \underline{F} is equal to the standard deviation of the equivalent normal distribution, and \underline{w}_f is assumed to be set at \underline{w}_2 , the angular frequency of the higher of the two possible signal frequencies \underline{w}_1 and \underline{w}_2 . (An angular frequency has units of radians/sec, and a value which is thus equal to the product of 2π and the frequency in Hz). The equivalent rectangular bandwidth of the Gaussian filter is $\sqrt{2\pi} \cdot F$ Hz. The decision model adopted is similar in all respects to that described above for the unbiased observer in a 2AFC discrimination task, where for two signals $\underline{x}_1(t)$ and $\underline{x}_2(t)$ having the angular frequencies \underline{w}_1 and \underline{w}_2 , and the detected energies \underline{E}_1' and \underline{E}_2' , the probability of the response $\underline{R}(\underline{x}_2 > \underline{x}_1)$ given that $\underline{w}_2 > \underline{w}_1$ is given by

$$p[R(\underline{x}_2 > \underline{x}_1)] = 1 - \Phi(y), \text{ where} \quad (3.13)$$

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y \exp[-(x^2/2)] dx, \quad (3.14)$$

$$y = \mu(\underline{E}_2' - \underline{E}_1') / \sigma(\underline{E}_2' - \underline{E}_1'), \quad (3.15)$$

$$\mu(\underline{E}_2' - \underline{E}_1') = \mu(\underline{E}_2') - \mu(\underline{E}_1'), \text{ and} \quad (3.16)$$

$$\sigma^2(\underline{E}_2' - \underline{E}_1') = \sigma^2(\underline{E}_1') + \sigma^2(\underline{E}_2') - 2r(\underline{E}_1', \underline{E}_2') \cdot \sigma(\underline{E}_1') \cdot \sigma(\underline{E}_2'). \quad (3.17)$$

Henning has shown that the energy difference $\underline{\Delta E}$, where $\underline{\Delta E} =$

$\underline{E}_2 - \underline{E}_1$ has the mean $\mu(\Delta E)$ given by

$$\mu(\Delta E) = E. [G(w_2) - G(w_1)], \quad (3.18)$$

and variance $\sigma^2(\Delta E)$ given by

$$\sigma^2(\Delta E) = 2N_o^2 \cdot \sqrt{(2\pi)} \cdot F \cdot T + 2N_o \cdot E[G(w_1) + G(w_2)], \text{ where} \quad (3.19)$$

$$G(w) = \{\exp[-1/2 (w_f - w)^2 / F^2]\}, \quad (3.20)$$

\underline{E} is the energy which would be detected from the two equal amplitude signals without the filter $H(w)$, and \underline{N}_o is the average power density as measured in 1 Hz bands of any random noise present in the pass band of the filter.

A Multiple-Band Energy Detector and the Effect of Frequency Uncertainty.

The same reasoning applies to a more elaborate multiple-band model (Green and Swets, 1966) for two complex harmonic signals each having \underline{k} components of equal amplitude, with a series of \underline{n} filters having centre frequencies at multiples of the fundamental frequency of the higher frequency complex signal, and filter widths \underline{F}_j . The observer's response may be assumed to be determined by the energy differences over all \underline{n} filters. The mean of $\underline{\Delta E}$ will be summed over the \underline{k} filters at which there is a signal, so that from equation 3.18

$$\mu(\Delta E) = \sum_{j=1}^{\underline{k}} E_j \cdot [G(w_{2j}) - G(w_{1j})]. \quad (3.21)$$

The variance of $\underline{\Delta E}$ will be summed over all \underline{n} filters, so that if the noise has a flat spectrum over the frequency band covered by the series of filters, from equation 3.19

$$\sigma^2(\Delta E) = \sum_{j=1}^{\underline{n}} 2N_o^2 \cdot \sqrt{(2\pi)} \cdot F_j \cdot T + \sum_{j=1}^{\underline{k}} 2N_o \cdot E [G(w_{1j}) + G(w_{2j})]. \quad (3.22)$$

An alternative to the multiple-band model is a single-band model (Green and Swets, *ibid*), which embodies the assumption that an observer can only monitor a single frequency-selective channel at any one time. A critical difference between these two models is that, according to the single-band model, the detectability of a signal should increase as components are added if and only if the additional components are detected in the same frequency-selective channel. The multiple-band model predicts that there should also be an improvement in detectability as components which are detected by other channels are added. Schafer and Gales (1949) found that the threshold for a two-component signal whose components were separated by more than the critical bandwidth was about 1.5 dB lower than the threshold of either component alone, and also reported a fall in threshold of up to 3 dB for complexes of 4 to 8 components. This result supports the multiple-band model. Gässler (1954) and Marill (1956) reported a decreased threshold only for complex signals whose components fell within one critical band, and not for signals with more widely separated components, which suggests the use of only a single critical band filter. Green (1958), Green, McKey, and Licklider (1959) and Hall and Sondhi (1977), however, have reported further results consistent with those obtained by Schafer and Gales, so that a multiple-band model is favoured by the bulk of such detection studies.

The multiple-band model also suggests an effect of frequency uncertainty, that is, if an observer knows which channel will detect an incoming signal, he may be able to restrict his attention to that channel, and thus limit the

variability of the detected energy by excluding noise contributed by other channels. Little is known about the effects of frequency uncertainty upon the frequency discrimination of human observers, but a number of studies (e.g. Tanner, Swets, and Green, 1956; Veniar, 1958; Creelman, 1960) have shown a detection decrement where the frequency of the test signal is uncertain. A further demonstration of frequency uncertainty effects upon detection has been provided by Johnson and Hafter (1980), who showed that if a pre-signal cue is given indicating which of several signal frequencies may be presented, the detectability of the cued signal increases.

The basic logic of the multiple-band model is not limited to the Energy Detection model, but can be applied to any model of auditory detection and discrimination which supposes that a number of frequency selective channels independently contribute statistical information about the signal. Such a generalized multiple-band theory also exhibits a pronounced similarity to the first stages of the spectral-pattern recognition models of residue pitch perception described by Goldstein (1973) and Wightman (1973b).

Predictions of the Energy Detector Model.

This model makes no prediction for the frequency dependence of the DL but such dependence could be introduced by making the filter width parameter F some suitable function of frequency. For the effect of signal level where N_0 is constant, $\mu(\Delta E)$ will be proportional to E' , and $\sigma(\Delta E)$ will be proportional to $\sqrt{E'}$, so

that the DL, which is proportional to the likelihood ratio $\underline{\mu(\Delta E)}/\underline{\sigma(\Delta E)}$, will vary with $1/\sqrt{E'}$. For a signal of constant amplitude and a varying noise level, $\underline{\mu(\Delta E)}$ is constant, and $\underline{\sigma(\Delta E)}$ will vary with a weighted sum of $\underline{N_o}$ and $\sqrt{\underline{N_o}}$, where the weighting is dependent upon \underline{F} , so that the DL should vary inversely with a weighted sum of $\underline{N_o}$ and $\sqrt{\underline{N_o}}$. For a signal of constant amplitude, with a constant noise level but a varying duration $\underline{T_s}$, since E' is proportional to $1/\underline{T_s}$, $\underline{\mu(\Delta E)}$ will vary with $1/\underline{T_s}$ and $\underline{\sigma(\Delta E)}$ with $1/\sqrt{\underline{T_s}}$, so that the DL should vary with $1/\sqrt{\underline{T_s}}$. This duration dependence can only be expected to hold where $\underline{T_s}$ is less than the integration time $\underline{T_I}$. A further limit to this prediction arises because the duration of a pulsed sinusoid affects its spectrum, such that as duration decreases, the spectral line at the frequency of the sinusoid broadens. This effect occurs because the spectrum contains components at frequencies differing from the frequency of the sinusoid by $1/\underline{T_s}$ Hz, and as $\underline{T_s}$ decreases, and $1/\underline{T_s}$ thus increases, the envelope of the spectrum covers a progressively greater frequency region. For this reason, at some duration primarily determined by the bandwidth of the filter, that portion of the energy of the signal passed by the filter $\underline{H(w)}$ will become significantly less than the total energy of the signal, and the rate at which the DL increases with decreasing duration will be greater than a rate proportional to $1/\sqrt{\underline{T_s}}$.

Temporal Theories of Frequency Discrimination.

Since the temporal coding of frequency in the firing pattern of the eighth nerve is not related to firing rate per

se, but is related in a complex way to the distribution of temporal intervals between successive nerve firings (Rose et al, 1967), temporal models of frequency discrimination involve the modelling of the distribution of such inter-pulse-intervals. Siebert (1970) has described a model based upon such a code, where it is assumed that the distribution of inter-pulse-intervals is a non-stationary Poisson process, and that the firing rate is a function of time, signal frequency and signal amplitude. Further assumptions based upon physiological observations are made to describe the total number of fibres, their spontaneous firing rates, their frequency selectivity, and the dependence of firing rate upon signal amplitude. Applying a complex mathematical argument which will not be described here, Siebert concludes that an optimally efficient observer employing the available temporal information should show a DL limited by the minimum

$$DL \geq 1/\sqrt{(3 \times 10^6 T_s^3/f + 1.5 \times 10^6 T_s \ln(A))}, \quad (3.25)$$

where f is the frequency in Hz of the sinusoidal signal, T_s its duration in sec, and A its amplitude in units of threshold amplitude. This model does not make any explicit predictions for the effect of signal-to-noise ratio, which may presumably be considered to have an effect upon threshold, and thus upon A. The predictions of interest here are that the DL will vary inversely with $\sqrt{\ln(A)}$, and inversely with $1/T_s^{3/2}$. Since the temporal code for frequency is tonotopically distributed in the auditory nerve fibers (Kiang, 1965; Rose et al, 1967), a multiple-band model (see above, p16) employing temporal

information is as plausible as a multiple-band Energy Detection model.

Tests of the Predictions of Theories of Frequency Discrimination.

Comparing these two theories with the available data for frequency discrimination, it is immediately apparent that the values of the DL predicted by Siebert's model are considerably smaller than those observed. For a 1000 Hz sinusoid at 50 dB above threshold and of 100 msec duration, Siebert's optimally efficient observer should show a DL of about 0.01 Hz, by comparison with observed DLs of 1 - 2 Hz (Weir, Jesteadt, and Green, 1977). Plainly, if real observers are using temporal information they are far from optimally efficient. The Energy Detector model, with an equivalent rectangular filter width of about 90 Hz at 1000 Hz will, as shown below, predict a DL very similar to that observed by Weir, Jesteadt, and Green. The predicted dependencies of the DL upon signal level show a similar form for both models over a range of signal levels from 5 to 30 dB above threshold, after which the Energy Detector model predicts a considerably faster rate of decrease for the DL as the signal level increases. The most comprehensive study of the effects of signal level (Weir, Jesteadt and Green, *ibid*), shows a function relating the DL to signal level which is very similar in form to that predicted by Siebert's model. The results of this study are illustrated in figure 3.3, together with the forms of the functions predicted by Siebert's temporal model and the Energy Detector model.

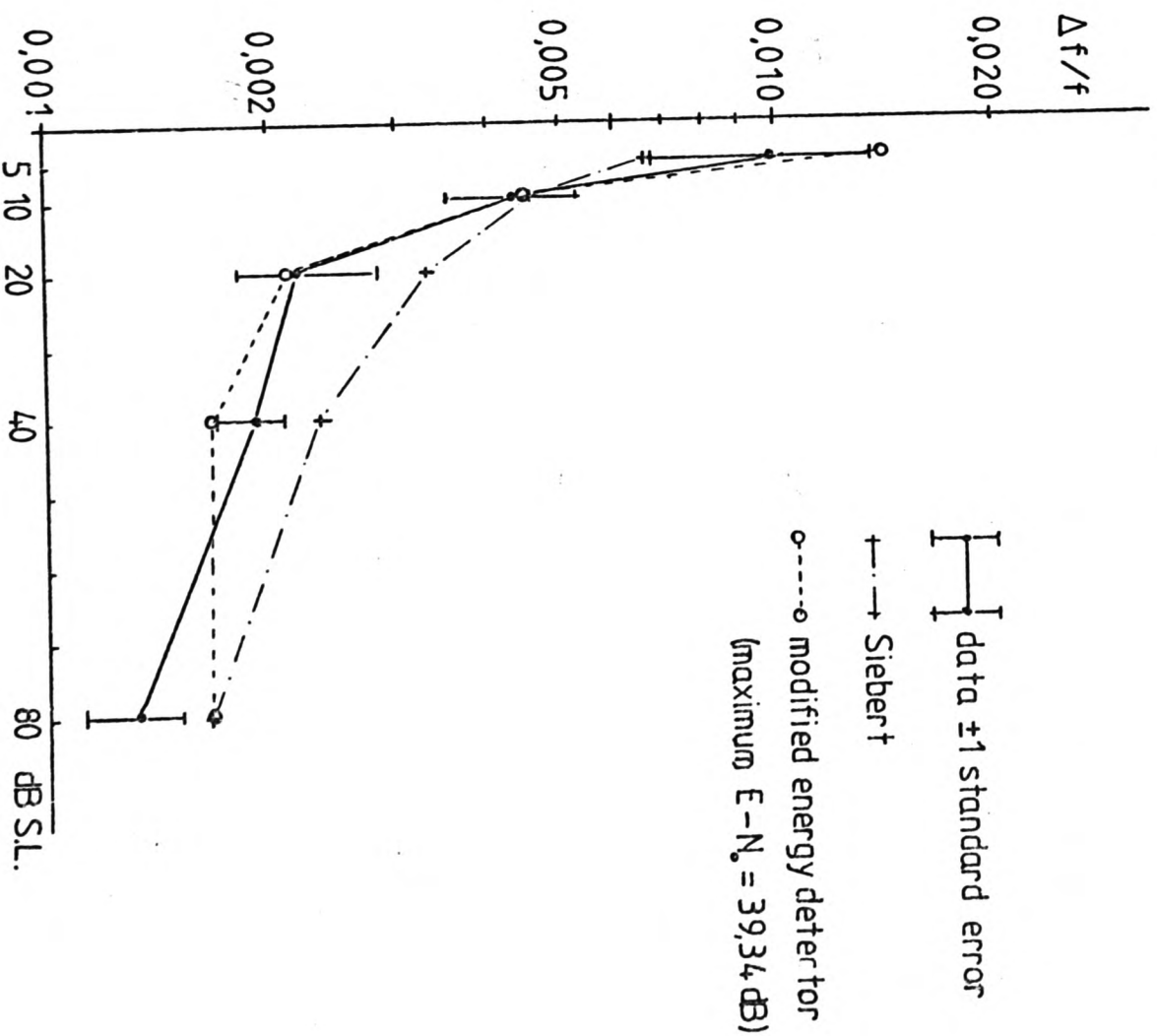
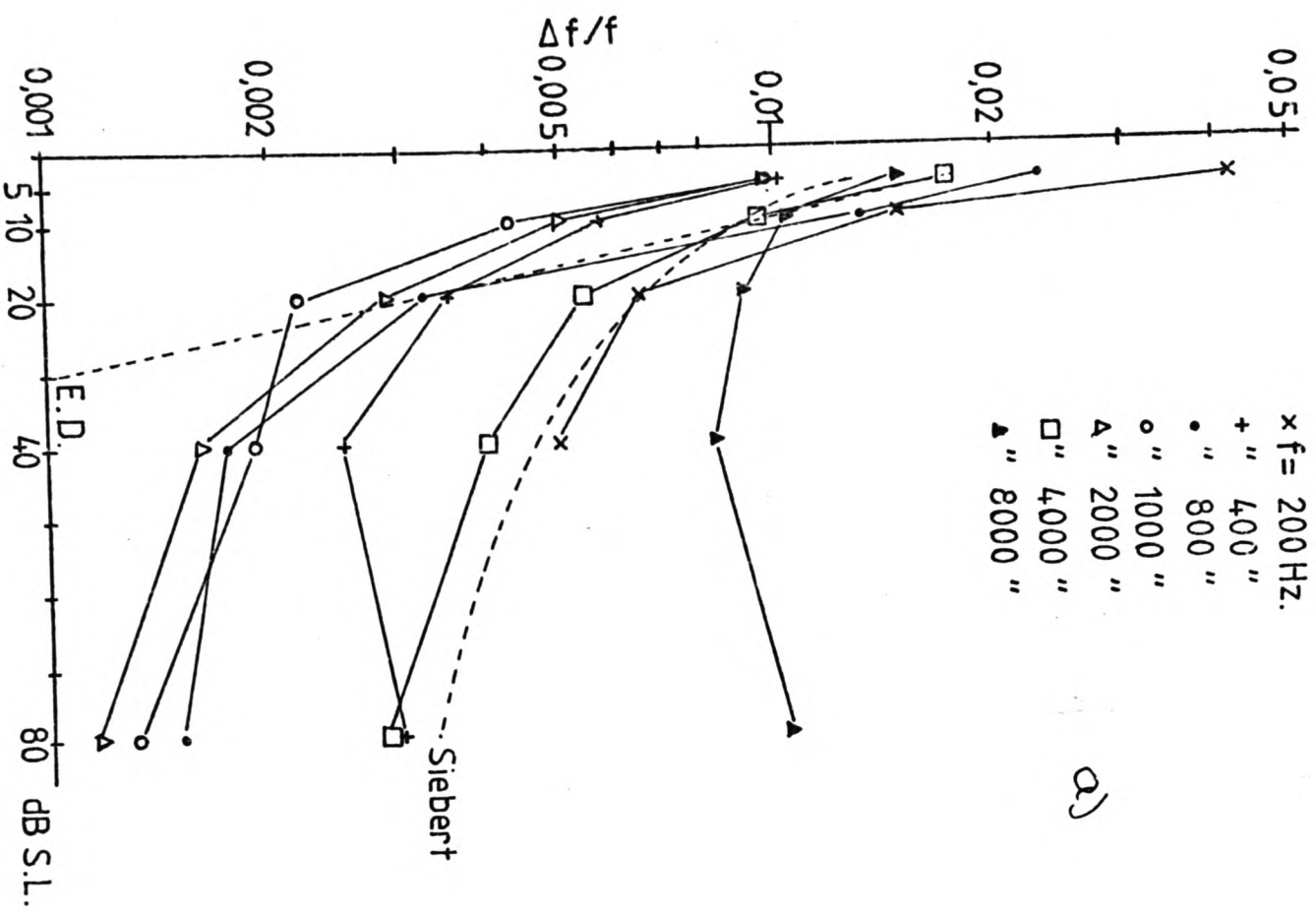


Figure 3.3 The effect of signal level on the difference-limen for the frequency of a sinusoid. Panel a) shows the DL's from Weir, Jesteadt and Green (1977) at frequencies from 200 to 8000 Hz expressed as $\Delta f/f$ where Δf is the frequency difference at which the observers were correct on 76% of trials in a 2AFC task. Also shown are the forms of the level dependencies predicted by the temporal model of Siebert (1970) and by a simple Energy Detector model (Henning, 1967a). Panel b) shows the results of Weir, Jesteadt and Green at 1000 Hz, together with the predicted DL's derived from a modified Energy Detector model having an internal noise process whose equivalent spectral density is 39 dB below the signal level. The form of the level dependence predicted by Siebert's model is also shown.

Henning (1967a) proposed a modification to the Energy Detector model which improves the fit of the model to performance at large signal-to-noise ratios. If the centre frequency ω_f of the Gaussian filter is taken to be a random variable, then the standard deviation of the energy difference ΔE includes a term which is proportional to the signal energy; the contribution of this term to the total standard deviation is small at low signal-to-noise-ratios, but at larger signal-to-noise ratios, this term is the major source of variability, and has the effect of limiting the minimum DL. An alternative modification of the Energy Detector model, which was proposed by Green and Swets (1966), and which also allows the model to account for the level dependence of the DL at larger signal-to-noise ratios, is to make the assumption that the likelihood ratio $\mu(\Delta E)/\sigma(\Delta E)$ has a minimum value determined by an internal noise process. This assumption is comparable to the assumption made by Siebert for the signal amplitude dependence of firing rate, and can be justified by reference to the observations of Rose et al (1971), that neither the firing rate nor the degree of phase locking to the signal shown by a single eighth nerve fibre changes once the signal is more than 40 to 50 dB above the fibres threshold, and that over the last 10 dB of that range, the effects of changes in signal level are small. The relation between the dynamic range of a fibre and the equivalent maximum signal power to noise spectral-density difference $E-N_0$ would depend upon the effective bandwidth of the fibre; for an equivalent rectangular bandwidth $\sqrt{2\pi}.F$ of 90 Hz, which is an appropriate value for this modification of the model and the data of Weir, Jesteadt, and Green at 1000 Hz, the

limiting value of $\underline{E-N}_0$ would be 19.5 dB above the dynamic range of the fibre. The DLs predicted by this modification of the Energy Detector model for a signal of 1000 Hz are included in figure 3.3. The model parameters giving these predictions were $\underline{F} = 36.20$ Hz, $\underline{T}_r = 0.298$ sec, and a maximum effective signal-to-noise ratio of 39.34 dB; for external signal-to-noise ratios of less than this limit, the value of $\underline{E-N}_0$ was primarily determined by the spectral-density of the background noise employed in this experiment, which was about 15 dB below the level of a 1000 Hz tone-burst of 100 msec duration and at 0 dB SL.

Suitable parameters of this modified Energy Detector model were found by calculating, for each of the signal levels in Weir, Jesteadt, and Green's experiment and for the frequency difference Δf corresponding to the mean DLs over four observers, the value of the predicted likelihood ratio. This ratio was compared with $1/\sqrt{2}$, which is the value of the likelihood ratio which should give performance of 76% correct in a 2AFC discrimination task, and the parameter values were adjusted by a Simplex minimization routine (Nelder and Mead, 1965; Note 2) to minimize the sum of the squared differences between the predicted ratio and $1/\sqrt{2}$.

Reasonably well-fitting parameters for the modified Energy Detector described by Henning (1967a) were also found, but the predicted DLs from this model do not fit Weir, Jesteadt, and Green's data at 1000 Hz as well as those predicted by the model with a limited maximum $\underline{E-N}_0$. The model parameters here were \underline{F}

= 33.14 Hz, \underline{T}_1 = 0.259 sec, and the variance of the centre frequency of the filter $\underline{H(w)}$ was 0.0991 Hz; the small value of this latter parameter suggests that it has very little effect upon the predictions. These parameters differ considerably from those given by Henning as the best-fitting parameters for his modification of the model and a similar experiment. Henning fitted the model with the assumption that the integration time \underline{T}_1 was equal to the signal duration, and found best-fitting values for \underline{F} and the variance of the filter centre frequency of 13.5 Hz and 9.5 Hz respectively for one observer, and 17.8 Hz and 10.15 Hz for a second. These filter width parameters correspond to an equivalent rectangular bandwidth of about 40 Hz, which is four times smaller than the typical critical bandwidth estimate. The greater filter widths and other parameters found here were obtained with the assumption that the integration time had some value not less than the signal duration and not greater than the sum of that duration and the inter-signal interval of 300 msec, and it appears that allowing a longer integration time leads to larger and more reasonable estimates of the filter width. The fitting procedure adopted here was able to reproduce Henning's best parameter values when the same constraint on the integration time parameter was imposed.

The critical bandwidth at 1000 Hz is typically 160 Hz (Scharf, 1970), which is somewhat greater than the equivalent rectangular bandwidth of 90 Hz found to fit Weir, Jesteadt, and Green's data at the same frequency. Houtgast (1974) has suggested that the conventionally estimated critical bandwidth is an overestimate of the effective bandwidth of auditory

frequency resolution, and has estimated from non-simultaneous masking studies a bandwidth of about 80 Hz ^{at} 1000 Hz, which is remarkably close to the best-fitting filter width for the Energy Detector model.

Few studies of the dependence of the frequency DL upon signal-to-noise ratio have been published, but that described by Henning (1967b) gives results in approximate accordance with the predictions of the modified Energy Detector models described above. Those data also show approximately the form predicted by Siebert's model if the parameter A in that model is equated with the difference between the signal-to-noise ratio as presented and that ratio at threshold, which for similar signals is typically about 16 dB (Green, 1976). As is the case with the effect of signal level, the actual values of the DL predicted by Siebert's model are two orders of magnitude smaller than those observed.

Several studies of the duration dependence of the DL are available (e.g. Henning, 1970; Liang and Chistovich, 1961; Moore, 1973a). The DLs obtained in these studies at the frequencies 200, 1000 and 4000 Hz as a function of the signal duration are shown in figure 3.4. The results of these three studies may all be described by the relation

$$\begin{aligned} \text{DL} &\approx 1/T_s, & T_s < t_1, \\ \text{DL} &\approx 1/\sqrt{T_s}, & t_1 < T_s < t_2, \\ \text{DL} &\text{constant}, & t_2 < T_s \end{aligned}$$

where T_s is the signal duration, and t₁ is frequency dependent. t₁ is equal to about 20 msec at 1000 Hz, to about 100 msec at

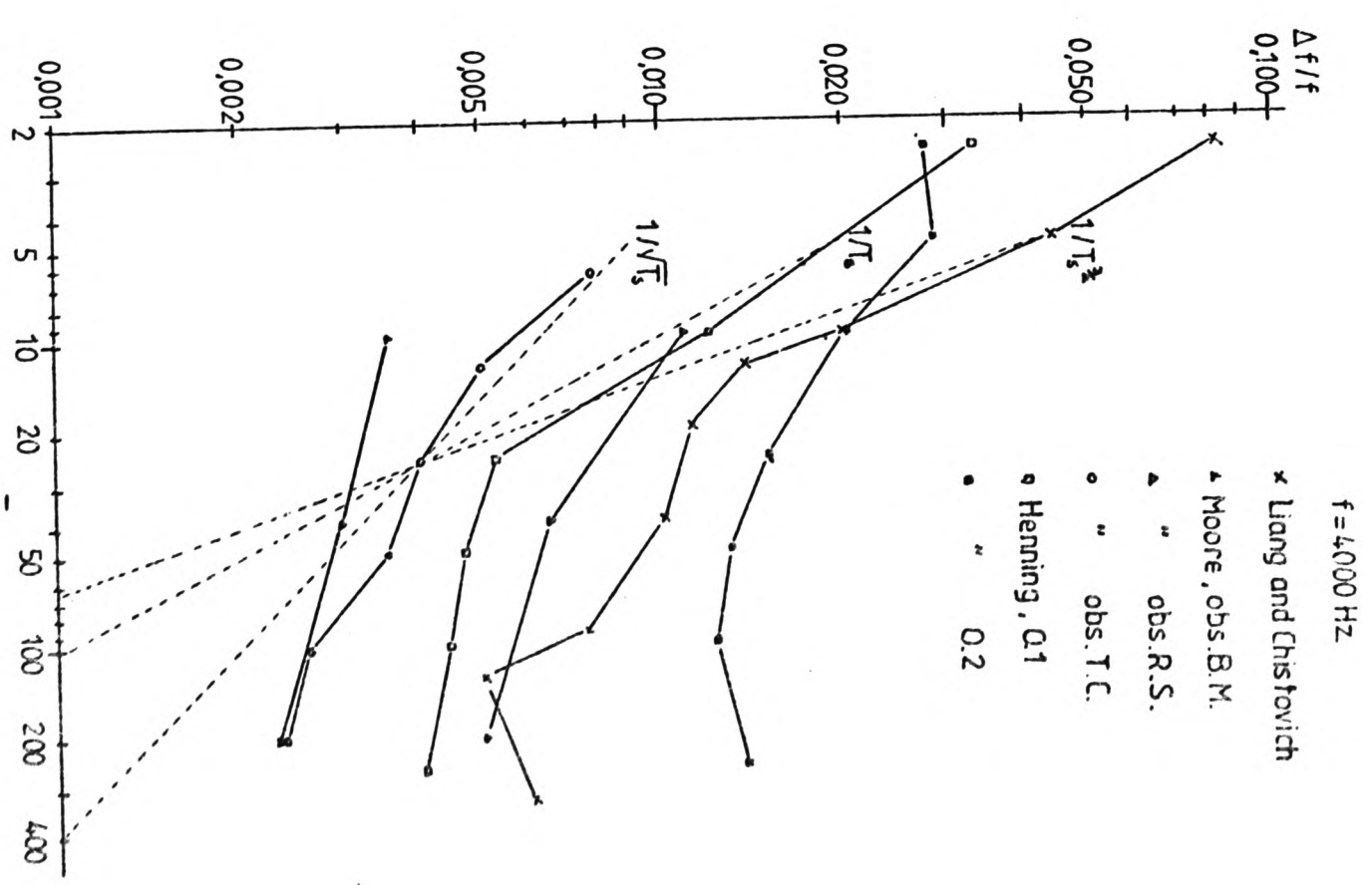
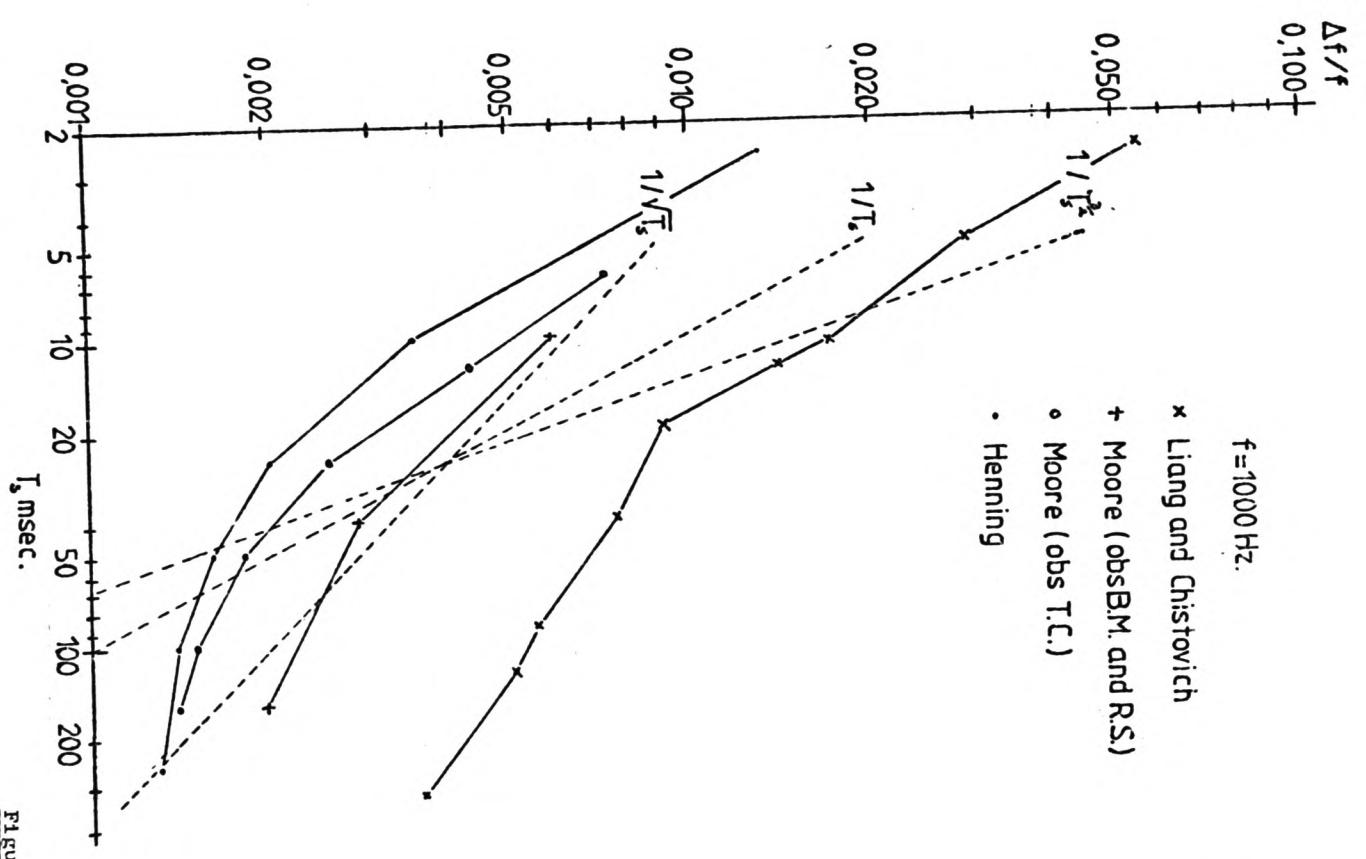
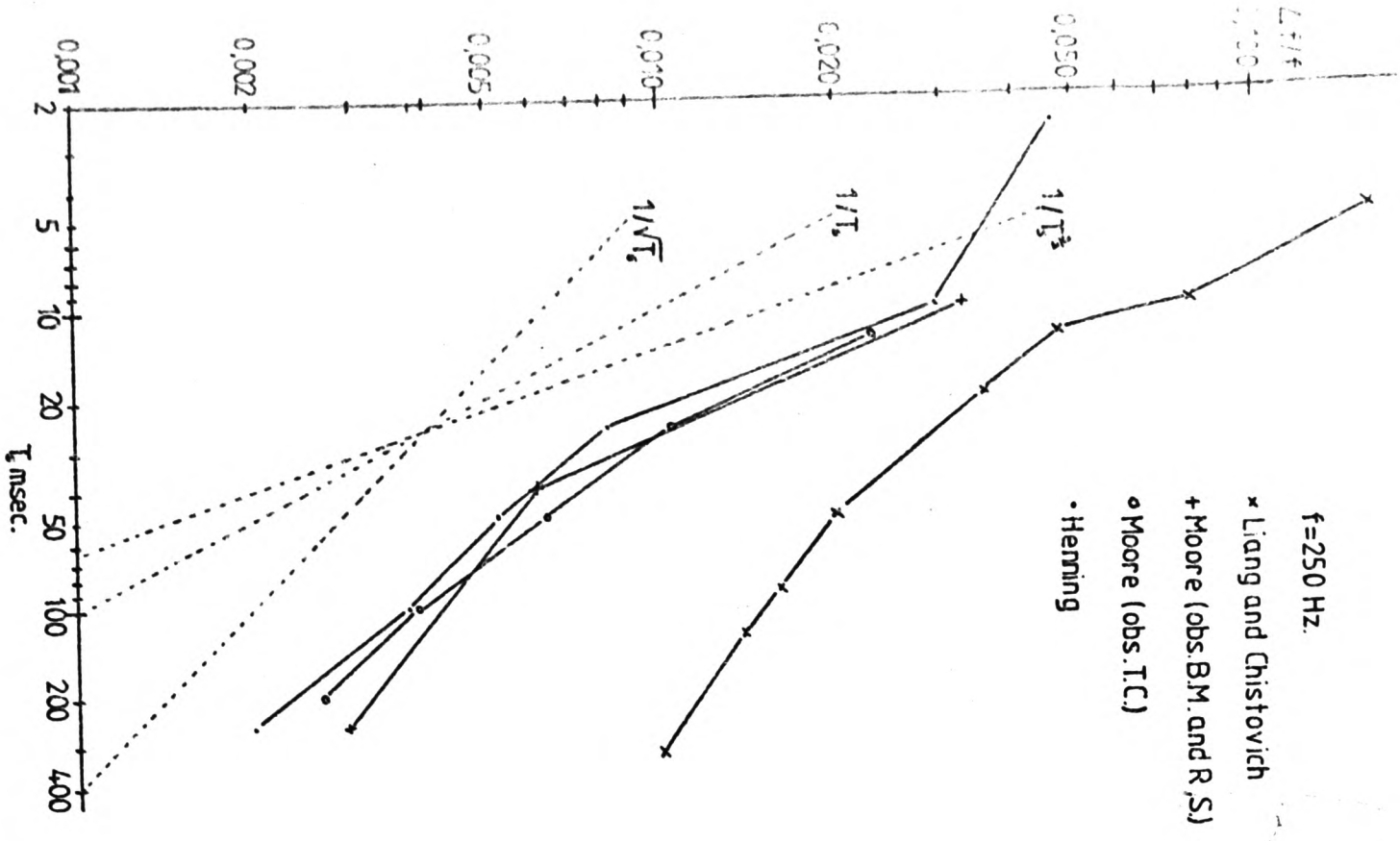


Figure 3.4 The effect of signal duration T_s on the frequency-difference limen $\Delta f/f$ for sinusoids at 250, 1000 and 4000 Hz. The data are taken from Liang and Chistovich (1961), Henning (1970), and Moore (1973a). The data from Liang and Chistovich is the mean of the performance of two observers, as is that of Henning except at 4000 Hz, where the performance of the two observers differed markedly; here the performance of each observer is shown separately. The procedure adopted by Liang and Chistovich differed from the 2AFC method employed by Henning and Moore, and this is the reason for the consistently larger Δf 's from the former study.

200 Hz according to the results of both Henning and Moore, but to about 40 msec according to those of Liang and Chistovich, and for frequencies of 4000 - 5000 Hz and above, $\underline{t_1}$ is 10 msec or less. Only Liang and Chistovich collected data for durations greater than 250 msec, and from their results, $\underline{t_2}$ appears to be independent of frequency and approximately equal to 150 msec. These studies confirm the prediction made from the Energy Detector model that the DL will vary with $1/\sqrt{T_s}$ for durations less than some constant integration time $\underline{T_I}$, where $\underline{T_I}$ may be equated with $\underline{t_2}$, and greater than some duration $\underline{t_1}$ at which the signal spectrum becomes broader than the filter bandwidth. The prediction from Siebert's model that the DL should vary inversely with $\underline{T_s}^{3/2}$ is clearly rejected for $\underline{T_s} > \underline{t_1}$, although it is difficult to reject such a relation for $\underline{T_s} < \underline{t_1}$, since only a small range of durations below $\underline{t_1}$ have been investigated.

Moore (1973a) has pointed out that at short durations, models based, as is the Energy Detector model, upon the transformation of frequency differences to amplitude or energy differences, must be modified to take account of the response of a resonant filter to the transient events which are inevitable consequences of the onset and offset of a finite duration signal. The response of a resonant filter to a transient event is an oscillation whose instantaneous frequency is equal to the centre frequency of the filter, and which decays at a rate inversely dependent on the bandwidth of the filter. The filter's response to a signal of finite duration will be the sum of this transient response and the response due to the steady-state part of the signal. Unless the signal frequency

and the filter's centre frequency are identical, this result will be a waveform whose amplitude and instantaneous frequency fluctuate. As the signal duration decreases, the response of the filter will become progressively dominated by the transient term, and thus, both the amplitude passed by the filter and the frequency of its output will become increasingly uncertain. Moore has shown that the uncertainty of the amplitude passed by such a filter implies that the product of the DL and the signal duration must be limited to some constant determined by the smallest amplitude difference which is discriminable by the decision process operating upon that difference. Moore assumed a discrimination acuity of 1 dB for a difference in amplitude, and, for rectangularly gated signals, argued that the DL resulting from a model of this type should be limited by a minimum value for the product of the DL and the signal duration of about 0.24. The assumption that amplitude differences smaller than 1 dB cannot be discriminated is consistent with studies of intensity discrimination, where for example, Jesteadt, Weir, and Green (1977) found intensity difference limens of 1 dB or more for signals up to 40 dB above threshold. The frequency discrimination data reported by Moore show instances at short signal durations and frequencies below 4000 Hz where the product of the DL and signal duration is as small as 0.05, and from these results he concludes that, at frequencies below 4000 Hz, frequency discrimination cannot be due to the use of energy or amplitude differences, but must be the result of the use of information provided by the temporal firing pattern.

There are several difficulties with this conclusion. In the first place, the signals used by Moore were not rectangularly gated, but had rise and fall times of 2 msec, so that his calculated limit for the product of the DL and signal duration is ^{slightly} higher than that appropriate to these signals. It is also possible that at ^{very} short signal durations, ^(≤ 10 msec) observers make deliberate use of cues provided by the peaks in the signal spectrum at frequencies differing from the steady-state signal by multiples of $1/T_s$; the use of such cues would serve to decrease the observed DL. A test of this possibility would be provided by an experiment similar to Moore's, but where the durations of the two pulsed sinusoids in a given trial differed. Here, the spectral components of the two pulsed sinusoids at frequencies other than their steady-state frequencies would be determined by both the steady-state frequency difference and the difference between the durations of the two signals, so that an observer would not be able to utilize these cues. Turnbull (1944), described such an experiment, where one pulsed sinusoid always had a duration of 500 msec, and the duration of a second sinusoid was varied. From the model of a 2AFC frequency discrimination task described above (pp 3-4-3-11), and Moore's (1973a) estimate of the minimum DL for a 2AFC task where both signals are of equal duration, the corresponding minima for signals of different duration may be calculated. Where there are two equal duration signals, the variances of the internal representations of frequency can be assumed to be equal, so that, from the model given above, the minimum standard deviation of each distribution $\sigma(f'_i)$ is approximately given by,

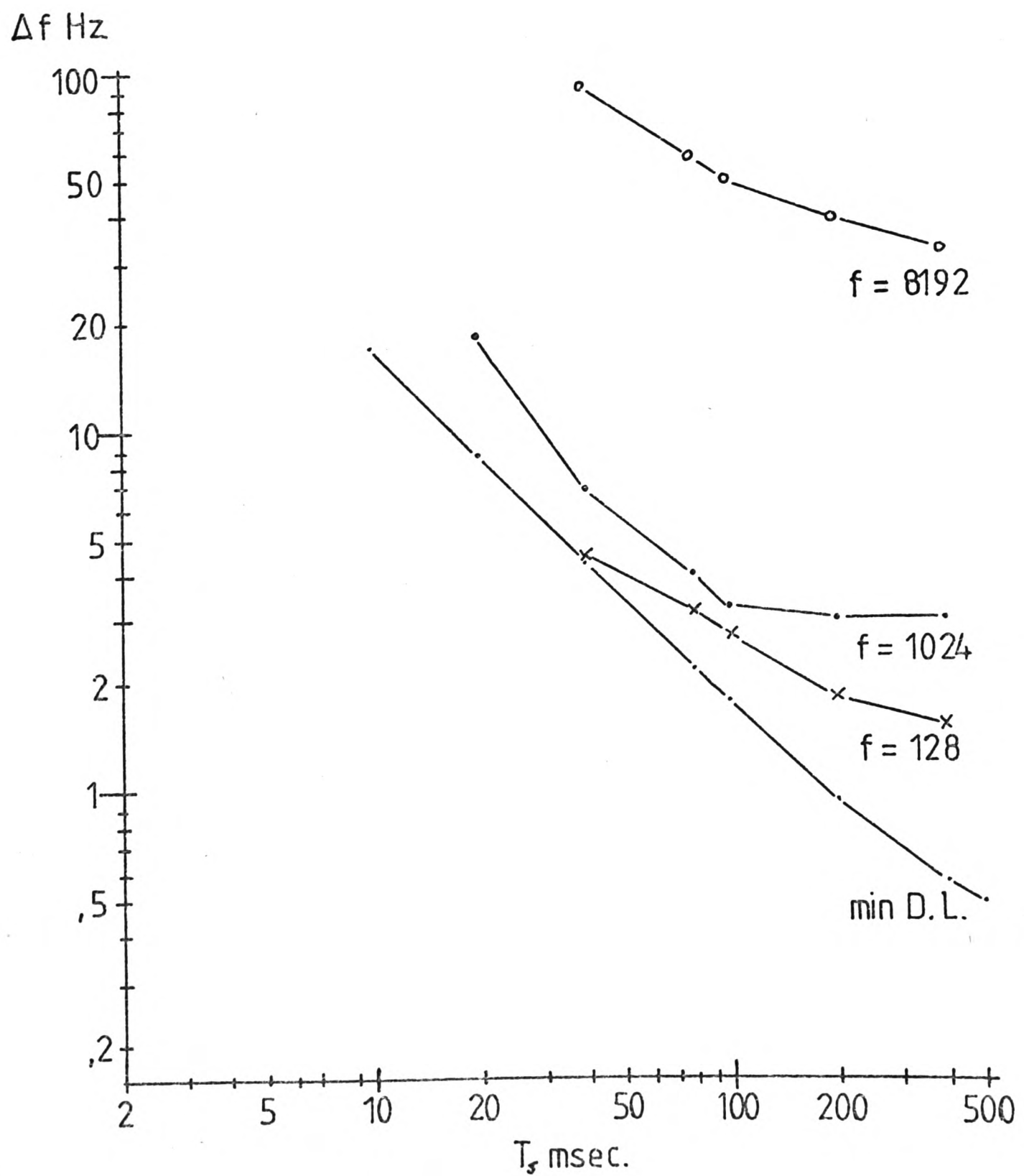
$$\min[\sigma(f'_i)] = \min(DL) = 0.24/T_s \text{ Hz.} \quad (3.26)$$

For the case of a comparison of the frequencies of sinusoids of differing durations, the assumption that the variances $\sigma^2(f'_1)$ and $\sigma^2(f'_2)$ are equal becomes unreasonable, and the expected DL must be found from

$$\begin{aligned} \min(DL) &= \sqrt{[\min(\sigma^2(f'_2)) + \min(\sigma^2(f'_1))]/2} \\ &= \sqrt{[(0.24/0.5)^2 + (0.24/T_2)^2]/2} \text{ Hz.} \end{aligned} \quad (3.27)$$

The minimum DLs predicted for this task with the constraints upon an amplitude or energy difference based model of frequency discrimination given by Moore have been calculated, and appear in figure 3.5 together with the results of Turnbull's study. Neither of Turnbull's observers showed a DL smaller than the predicted minimum at any duration or at any of the three signal frequencies used in the study, and, except at 8192 Hz, where the observed DLs are rather large, the agreement between the data and the predicted minimum DLs is remarkably close. While this comparison is based upon a study which was carried out with equipment which doubtless would not compare with the precision of modern devices, the experimental method was comparable to that used by Moore (1973a). However, the smallest DL found by Turnbull at 1000 Hz was about twice those found by Moore for comparable durations, and Turnbull was unable to measure any DLs for durations of less than 20 msec. A replication of the study would be required to settle the issue, but Turnbull's data are consistent with the hypothesis that cues derived from the transient part of the spectra of sinusoidal tone-bursts are responsible for Moore's (1973a) finding that the DL for short

Figure 3.5 The DL for a sinusoid estimated by Turnbull (1944), where the first of two tone-bursts was always 500 msec in duration and the second was of duration T_s . The figure also shows the minimum value of the DL as calculated from the limit derived by Moore (1973b): see text for description of the calculations.



duration sinusoids is not limited by the ability of observers to make amplitude discrimination judgements at the steady-state frequency of these signals.

A further, and perhaps more fundamental difficulty with Moore's criticism of the models based upon the detection of an amplitude or power difference is that if the amplitude passed by a filter responding to a short duration signal is uncertain, the frequency of the response of the filter must also be uncertain, so that a frequency coding based upon the temporal pattern of neural activity will be subject to a similar limit. It is possible of course, that the auditory system is inherently better at the temporal coding of frequency than at coding amplitude or energy, and Siebert's temporal model, which is based upon observations of neural responses, suggests that this is the case, although Siebert's model greatly overestimates the precision of frequency coding.

Modifications of Siebert's model.

Luce and Green (1974) have presented a modification of Siebert's (1970) model, which assumes a less than optimally efficient observer, and predicts more reasonable values for the frequency DL. They do not however develop any different predictions for the effect of signal duration. Goldstein and Srulovicz (1977) have presented another modification of Siebert's model, which is able to predict reasonable values for the DL and also the dependence of the DL on $1/\sqrt{T_s}$, and with appropriate parameter values, can account for Moore's (1973a)

data. This model does not not, however, make direct predictions for the effects of signal and noise levels; the terms in the model related to these parameters represent the firing rate of the auditory fibres and the degree of synchrony of firing to the signal. Since Goldstein and Srulovicz do not make explicit the effects of signal and noise levels upon these two factors, and the available physiological data is sparse, particularly with respect to the effects of noise, this model is of limited usefulness here.

The Signal Detection Model and Temporal Theories of Frequency Discrimination.

The stochastic rate functions which have been proposed for the temporal coding of frequency suggest that the representation of frequency in a single auditory fibre is not a normally distributed random variable. However, since a large number of fibres will each exhibit approximately independent representations of the frequency of a signal, the sum of these representations may, by the Central Limit Theorem (Hays, 1963), be considered to be normally distributed. For this reason, the decision model described above remains appropriate for the analysis and description of human frequency discrimination performance regardless of the form in which the internal representation of frequency is coded. If it is the case that frequency is coded temporally, and if it is also the case that the frequencies of the several components of a complex signal are coded independently in what may be thought of as independent neural channels or ensembles of auditory fibres, then this

decision model may be elaborated to describe frequency discrimination for the several components of the complex in a manner directly analogous to the elaboration of the Energy Detector model described above. In that model, the information upon which a decision is based is the difference between the energies at the output of a band-pass filter to two signals. For a model based upon the use of a temporal code for frequency, the corresponding information will simply be the difference between the neural estimates of the frequency of the two signals.

Summary.

1/ The application of the Theory of Signal Detectability (Green and Swets, 1966) to a 2AFC frequency or pitch discrimination task has been described. This theory provides a statistical method whereby the parameters of a normally distributed internal representation of some stimulus may be estimated. This method will be applied below to the estimation of the theoretical parameters $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$, the relative standard deviations of internal representations of the frequency or pitch of a component of a complex signal, and the residue pitch of a complex signal.

2/ Several theories of frequency discrimination, all consistent with the above decision model have been described. Both spatially and temporally based theories were considered. Although the richness of the available neural information strongly suggests that temporal coding plays an important role

in the coding of frequencies below about 4000 Hz, the available data are not sufficient for the rejection of the spatially based Energy Detector model, which is able to account for all but one aspect of that data: The Energy Detector appears unable to account for frequency discrimination for short duration signals (Moore 1973a), but short duration signals produce spectral cues remote from the signal frequency, and if these are taken into account, the Energy Detector model may be more successful.

3/ There is at least one experimental comparison which may permit the rejection of the Energy Detector model. According to that model, the detected energy and the signal duration are directly inversely related, so that at least over the range of durations between the time-constant of the theoretical filter and the integration time of the detector, signal duration can be traded for signal power, and the DLs observed for signals presented at equivalent signal energy to noise spectral-density ratios should be the same. The prediction of Goldstein and Srulovicz's (1977) model that the DL should depend inversely upon the square root of signal duration is derived quite independently of the model's predictions about signal and noise level effects, so that this model does not imply such a trading relationship. Two pairs of experiments whose results may be compared with this prediction, where the DL measured was the DL of the residue pitch in the region of the fundamental frequency of a two-component harmonic signal, as a function of the signal-to-noise ratio (Experiments 8 and 10), and as a function of signal duration with a constant noise spectral-density (Experiments 12 and 13), are described in Chapters 7 and 8

respectively.

CHAPTER 4

Abstract.

In this chapter, reasons are given for the choice of the stimuli employed throughout the experimental studies described in this and subsequent chapters. The chapter also contains a description of the apparatus employed for all of the experiments described below. In addition, the methods by which the first two of these experiments were carried out are described, and this description also serves as a general introduction to the techniques employed in the majority of the other experiments described below, these being Experiments 3 to 13, which are described in Chapters 5 to 8, and which all make use of the 2AFC frequency or pitch discrimination task described in the previous chapter.

The two experiments described in this chapter are studies of the effects of presentation level upon discrimination for the residue pitch of a two-component harmonic signal and frequency discrimination for a sinusoid at the frequency of one of the components of one such a complex. The results of the first experiment indicate certain difficulties in the application of the 2AFC pitch discrimination task and the associated decision

model to the chosen stimuli. These difficulties are discussed, leading to alternative stimulus configurations which have been employed in the pitch discrimination experiments described in subsequent chapters.

The Choice of Stimulus Parameters.

In order to examine the parameters of spectral-pattern recognition models of the perception of residue pitch and their inter-relations, it is clearly necessary to choose fundamental frequencies and harmonic numbers that lie within the existence region for residue pitch. A test of the contrasting predictions of spectral-pattern and periodicity based theories of residue pitch perception derived in Chapter 2 also requires that the components of the signals should be well-resolved. In addition, it is necessary to ensure that the spectra of the signals employed can be well controlled, and in particular, that no additional components are introduced by auditory non-linearities. A further desirable property of the signals employed is that they contain the minimum amount of redundant information, because a test of the predictions made in Chapter 2 will be valid only if the observers are making full use of all of the available stimulus information. For the latter reason, the signals employed in the experimental studies described below contain only two components, two being the minimum number of components required for the definition of the fundamental frequency of a harmonic signal. As the predictions are concerned with the residue pitch of harmonic signals, the signals employed here are harmonic.

A considerable number of studies of residue pitch have investigated signals whose fundamental frequencies lie at about 200 Hz, and since this fundamental frequency is commonly employed and known to be conducive to the perception of residue

pitch, the signals employed here also have a fundamental frequency in this region. The limits of the existence region of the residue pitch for two-component harmonic signals can be derived from the results of at least three studies, where two of these are concerned only with signals where both components were presented to both of the observer's ears (Smootenburg 1970, Sutton and Williams 1970), and the third (Houtsma and Goldstein 1972) also considered signals whose two components were dichotically presented. The existence region derived from these studies, for conditions where the influence of auditory distortion products was eliminated, has, for a fundamental frequency of 200 Hz, an upper limit for the lower component at about the 8th harmonic. The 2nd, 4th and 8th harmonics have pitches which are one, two and three octaves respectively above the pitch at the fundamental frequency, and are thus closely related in a musical sense to the pitch at the fundamental frequency. For this reason, it might be argued that the presence of one of these components is in itself a sufficient cue for the pitch judgements which are to be proposed as evidence for the perception of a residue pitch at the fundamental frequency, and it is desirable to avoid as far as possible complexes containing these components.

For the above reasons, the following three signals have been employed as the complex signals whose residue pitch was to be judged. These signals were always of the form

$$f(t) = a \sin(2\pi n.f.t) + a \sin(2\pi (n+1)f.t),$$

where a is constant for each component, and n may have the value 4, 5 or 6. Although a signal containing the 4th harmonic is

included in this set, at least one of the experiments employing each of the selected independent variables allows an independent measure of the judged pitch of this and of another signal, so that the extent to which a cue derived solely from the 4th harmonic is responsible for the observed pitch judgements can be ascertained from the data. For all of these signals, the frequency separation between adjacent components is equal to the fundamental frequency, 200 Hz, which is never greater than Scharf's (1970) estimates of the critical bandwidth at these component frequencies; it is thus reasonable to assume that the two components of each signal may be separately resolved, and hence, to assume that their internal representations are approximately independent.

The Experiments to be Described.

The general strategy adopted for the series of experiments described here has already been introduced, and in summary is as follows. First, the parameter $\sigma(f'_j)/f_j$, the relative standard deviation of the internal representation of the j th component of the complex signal, is estimated as a function of the presentation conditions of the two components (see Experiment 3, Chapter 5, and Experiment 7, Chapter 7). From these values, the theoretical probability density function of the pitch values of the signal under the same presentation conditions is predicted from the spectral-pattern recognition models described above. A minimum value for one parameter of a comparable density function, $\sigma(f'_0)/f_0$, the relative standard deviation of that mode of the density function whose mean is at

the true fundamental frequency of the signal, was also predicted from a periodicity theory.

The predicted values of $\sigma(f'_0)/f_0$ are then compared with an estimate of that parameter from a pitch discrimination task for the residue pitch of the signal (see Experiments 5 and 6, Chapter 6, and Experiments 8, 9 and 10, Chapter 7). The parameter of the signal presentations which is varied in Experiments 3 to 6 is the sensation-level of the two components of the signal. In Experiments 7, 8, 9, and 10 the signal-to-noise ratio is varied. Experiments 12 and 13 are intended to estimate $\sigma(f'_0)/f_0$ as a function of signal duration. Finally, the areas of the various modes of the observed pitch distributions are estimated as a function of signal-to-noise ratio and signal duration from a pitch comparison task in Experiments 14 and 15 respectively, and compared with the theoretical pdfs predicted from the spectral-pattern recognition models with $\sigma(f'_j)/f_j$ estimated from the discrimination experiments. These comparisons are all made within the performance of the same two observers, and for conditions of stimulus presentation which are, as far as is possible, comparable in each of the experiments where the same independent variable is manipulated. The first two experiments described below are exploratory in nature, and rather than forming a part of the above strategy, they are primarily concerned with establishing that with the available observers and experimental methods, the subsequent stages are feasible.

Apparatus for the Experiments.

All of the experiments to be described were carried out with the same or a very similar configuration of equipment. The signals were generated with a digital computer (DEC PDP 8/e) equipped with two 10-bit digital-to-analogue converters. All of the signals were harmonic, and their components were added in sine phase. The signals were low pass filtered at 1.5 KHz with a slope of 48 dB/octave (Kemo VBF/3), ramped on and off by a computer controlled electronic switch giving an approximately exponential rise and fall (Grason-Stadler 829E), and stored on an analogue tape recorder (Revox A77, Half-track, High-Speed) loaded with high-quality tape (Agfa PEM 268/368). The signals were generated at half of the required frequencies, and the stimuli were played-back from the tape at twice the speed at which they were recorded, the playback speed being 15 inches/sec. The effective sampling rate of the signals at playback was approximately 16 KHz. Measurements of the speed stability of the tape-recorder are presented in table 4.1. The equipment employed for the generation and presentation of the signals is illustrated in figure 4.1. Signals controlling the warning lamp and the gating by the electronic switch of the background noise were derived from the tape-recorded stimuli by means of the dia-pilot. The signals were presented monaurally through earphones (Rogers Ravensbrook, with supra-aural cushions) to the observer's left ear, and the observer was seated within a sound-insulated booth (Amplivox). The voltmeter from a sound-level meter (Dawe 1400G) was used for day-to-day calibration of signal and noise levels; since the frequency

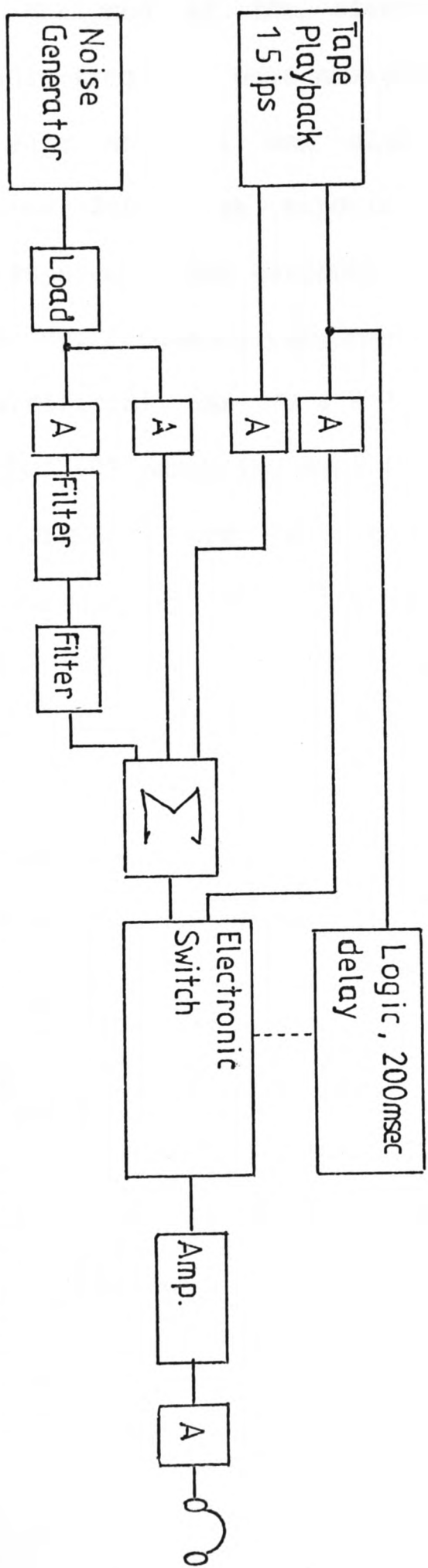
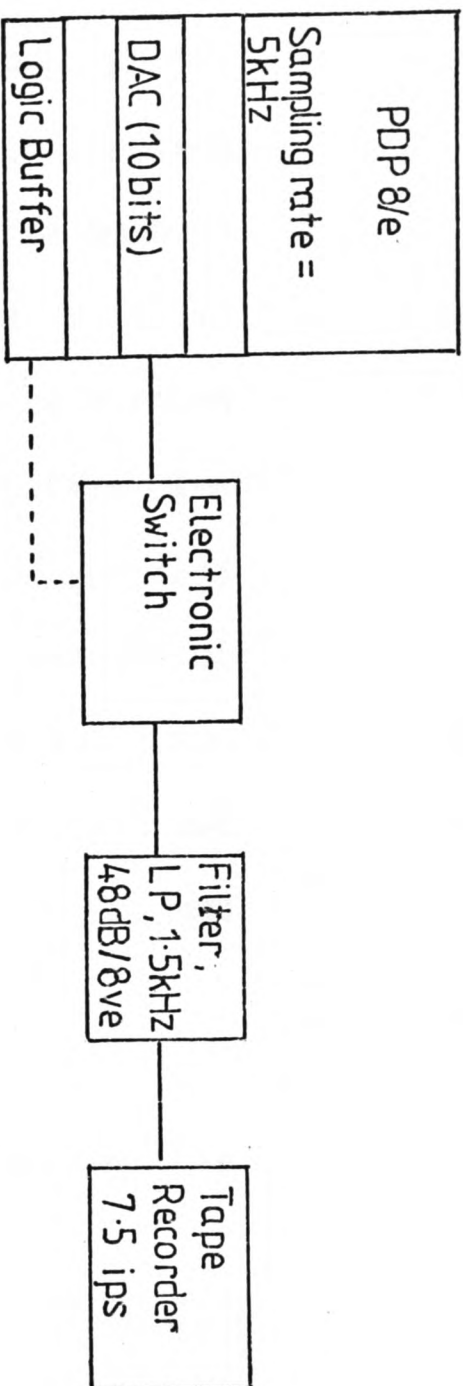


Figure 4.1 Block diagram of equipment used for the experiments. The upper half of the figure shows the equipment for recording, and the lower half that for playback. Solid lines represent signal paths. Dotted lines represent control logic paths.

response of this device was unknown, each stimulus tape included a reference level tone, which was always of the same frequency (1047 Hz). The results of spectral analyses of the signals used for the experiments described below are presented in table 4.2. The signals were analysed at the electrical input to the earphone using a audio-frequency wave analyser (Hewlett Packard HP 3580A). The wave analyser was also employed to measure signal-to-noise ratios for those experiments where a noise background was presented. The earphone itself was calibrated for amplitude linearity, frequency response and sensitivity with the aid of an artificial ear (B & K 4153, with a type 4134 microphone and a type 2607 measuring amplifier); the results of the linearity measurements are presented in table 4.3. The measured frequency response of the entire system including the earphone was flat within ± 2 dB between 200 and 2500 Hz, and rolled off at about 6dB/octave below 200 Hz.

The observer responded by writing a code representing his judgement on a duplicated response sheet.

Table 4.1. Speed stability of Revox A77 tape-recorder. Measured with a Ferrograph test set, type RTS2. All measures are peak readings for over a record-replay cycle.		
Tape Speed	Drift	Wow and flutter.
7.5 ips	<0.05%	Start of 7" reel , <0.07%
		Centre of 7" reel, <0.07%
		End of 7" reel , <0.15%
15 ips	<0.05%	Start of 7" reel, <0.07%
		Centre of 7" reel, <0.07%
		End of 7" reel, <0.09%

Table 4.2. Wave analysis of signals. The signals were analysed at the electrical input to the earphone after tape-recording and playback. Two-component harmonic signal, $f_0 = 209.4$ Hz, $f_1 = n.f_0$, $f_2 = (n+1)f_0$. The entry is the level in dB at f relative to the level at f_1 .									
n	Level at f_1		level at f_2		level at $2f_1 - f_2$			level at $f_2 - f_1$	
2	0		0		< -50			$f_2 - f_1 = 2f_1 - f_2$	
4	0		0		< -50			-48	
5	0		0		< -50			-48	
6	0		0		< -50			-48	
10	0		0		-45			< -50	
Five component signal, as reference signal in Experiments 4, 8, 9, 11, 13 and 14. $f_0 = 209.4$ Hz, components added at equal amplitude at the 1st, 2nd, 7th, 8th and 9th harmonics. The entry is the level in dB at $n.f_0$ relative to that at 209.4 Hz.									
n =	1	2	3	4	5	6	7	8	9
	0	+2	-45	-40	-38	-37	-1	-1	-1
Apart from a 50 Hz mains ripple, and harmonics of that frequency, all of which were at least -45 dB relative to the signal components, only the components in the table were measurable. These measurements were made using a Hewlett Packard HP 3580A wave analyser set to a 10 Hz bandwidth.									
Table 4.3. Calibration of earphones (Rogers Ravensbrook). Amplitude linearity. The tabulated input level is arbitrarily scaled, and the output level is given in dB SPL. The input signal was a 1000 Hz sinusoid.									
Input level (dB)					Output level (dB SPL)				
0					53.5				
10					63.5				
20					73.5				
30					83.5				
40					94.0				
Table 4.4. Estimated thresholds of the three observers at 1047 Hz in quiet. Thresholds were determined by a 2IFC detection task.									
Observer		Threshold, dB SPL.							
A.F.		12							
G.S.		8							
B.P.		6							

General properties of the stimuli.

In a number of respects, the stimuli for the experiments described below were similar. For the sake of brevity, their typically common properties are described here. Except where the details given here are incorrect, or of particular importance, the stimulus descriptions in the following chapters will omit some of this information.

The average fundamental frequency f_0 of the signals was 209.4 Hz, a value dictated by the sampling rate of the digitally generated signals and the available computer memory; in the interests of simplicity, this value will on occasion, be referred to nominally as 200 Hz. In all the residue pitch discrimination experiments, the fundamental frequency difference Δf between the two signals was small by comparison with the average fundamental frequency, and also by comparison with the frequency difference between the various residue pitch values which the signals would be expected to exhibit as determined by any of the theories of the residue pitch described above.

The temporal properties of the stimuli for Experiments 1 to 11 were always the following. The signals were gated on and off with an exponential envelope and a rise and fall time of 12.5 msec, so that no clicks were audible at their onset or offset. The duration of the steady-state portion of each signal was 87.5 msec, and the half-power duration of each was 100 msec. For all of the experiments, the interval between the offset of the first and the onset of the second signal in a trial was 200 msec, which was considered sufficient to ensure the absence of any peripheral interaction or masking between the two. The intertrial interval was 2850 msec, which allowed sufficient time for the observer to make his written response.

The Purpose of Experiments 1 and 2.

The first of these experiments was intended to serve several purposes; the experiment was preliminary to those

subsequently described in that it constituted an examination of the feasibility of the projected investigations of the statistical properties of the residue pitch of two-component harmonic signals. The experiment also provided the observers with experience of a psychophysical task requiring judgements of that pitch. The experimental task was so constructed that the observers' judgements could provide clear evidence of each observer's ability to perceive a residue pitch from these rather simple signals. Such a demonstration has some importance, as Evans (1978) has claimed that the residue pitch of certain two-component complexes is not highly salient. Finally, the experiment served a theoretical purpose as an investigation of the effects of the presentation level of the signal upon discrimination of the residue pitch of the signal, while the second experiment investigated the effects of this same independent variable upon frequency discrimination for one of the sinusoidal components of the complex signal. The effect of frequency upon the DL for a sinusoid is small over the range 800 to 1400 Hz (Shower and Biddulph, 1931; Harris, 1952), and the DL for a sinusoid at 1256 Hz, the 6th harmonic of the fundamental frequency used here, was taken to be representative of the DL for all components of these complexes from the 4th to the 7th harmonic.

Experiment 1.

Observers.

Three observers took part in the experiment. One of these

was the author; the remaining two were paid for their services. None of the observers suffered from any known hearing defect. All three were experienced amateur musicians, but without any extensive formal musical training.

Stimuli.

The task presented to the observers was that of comparing the residue pitches of two harmonic complexes. In order to ensure that cues derived from the frequencies of single components could not lead to better than chance performance, the two signals presented on a given trial never contained the same harmonic of their respective fundamental frequencies, and the direction of the change of the component frequencies within a trial was made independent of the change in fundamental frequency. The signals were of the form

$$\begin{aligned} \text{a)} \quad f(t) &= a \sin(2\pi n(f_0 + \Delta f / 2)t) + \\ &\quad a \sin(2\pi(n+1)(f_0 + \Delta f / 2)t), \text{ or} \\ \text{b)} \quad f(t) &= a \sin(2\pi n(f_0 - \Delta f / 2)t) + \\ &\quad a \sin(2\pi(n+1)(f_0 - \Delta f / 2)t), \end{aligned}$$

where a was equal for each component, and n was equal to either 4 or 6. Because Δf was small with respect to f, Δf was also small with respect to the frequency difference between the higher component of the signal where n=4 and the lower component of the signal with n=6. On a given trial, one of two pairs of signals were presented. Each pair contained a type a) signal and a type b) signal, and n always differed between the two signals. These two signal pairs were presented in each of the two possible temporal orders. Given these four possible ordered

pairs, which were presented equally often, if the observer were to judge the relative pitches of the two signals not on the basis of a residue pitch, but on the basis of frequency differences between the components of the two signals, he would only be correct in terms of the relative fundamental frequencies of the signals on 50% of the trials, while judgements made from the residue pitches of these signals should lead to a performance of between 50% and 100% correct.

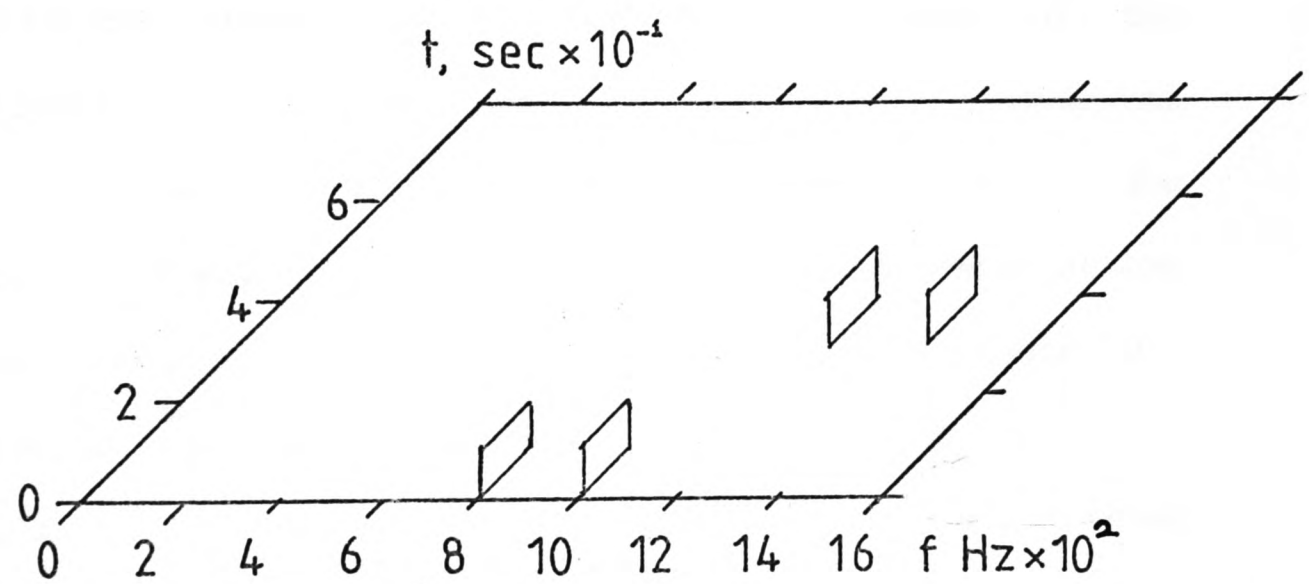
The amplitudes of the two components were always equal to one another. This amplitude was an independent variable, and was such that the level of the 5th harmonic, this being at a nominal 1000 Hz, was either 15, 20, 30, 40 or 50 dB above the observer's threshold with a quiet background. Thresholds were determined individually for each of the observers by a Two-Interval-Forced-Choice detection method. Thresholds for the three observers are given in table 4.4. These signal levels are hereafter described as sensation levels (SL); although this description may not be strictly correct for components at frequencies other than 1000 Hz, it is less cumbersome than more correct alternatives, and in any case, the true sensation levels at frequencies between 800 and 1400 Hz should not differ greatly (Robinson and Dadson, 1956).

The time-course of a trial is illustrated on figure 4.2.

Procedure

Trials were presented in blocks of 100, during which $|\Delta f|$

Figure 4.2 Schematic representation of a single trial from Experiment 1. The two signals occurred in both possible orders.



remained constant, and each of the four possible ordered pairs occurred 25 times in a pseudo-random order. The observer was instructed to compare the pitch of the second signal with that of the first, and to respond by marking a response sheet with the letter "H" if he considered the pitch of the second signal to be the higher, and the letter "L" otherwise. Observers were asked to try to ignore the pitches of the individual components of the signals, and to concentrate upon a lower pitch resulting from the combination of each signal's components. Prior to each test session, the observers were given a sequence of 40 signal pairs for which $|\Delta f|$ was approximately 24 Hz; each pair was followed by the illumination of a lamp, which lit once if the first signal of the pair was that of higher fundamental frequency, and twice if the second signal was the higher. The first 20 of these pairs also contained a component at the fundamental frequency of the signal. No immediate feedback was given during the remaining trials, but each observer was informed of his general performance level and of any marked inconsistencies in his performance when he next attended.

Each test session, approximately one hour in duration, consisted of eight blocks of 100 trials. The absolute value of Δf was constant for each block, and had maximum and minimum values of $0.03248 \cdot f_0$ and $0.00203 \cdot f_0$ respectively. For a given signal level, the blocks of trials were presented equally often with $|\Delta f|$ increasing and decreasing. The observers each received at least 100 trials at each combination of signal level and $|\Delta f|$. Two stimulus tapes were prepared, each with the stimuli recorded in a different pseudo-random order, and each

observer received each tape equally often; the two orderings of the change of $|\Delta f|$ over blocks were obtained by playing the tapes in both forward and reverse directions.

Experiment 2.

The three observers were the same as those who took part in Experiment 1. The signal presented in each of the two observation intervals was a single sinusoid, at a frequency of $6 \times (209.4 + \Delta f/2)$ or $6 \times (209.4 - \Delta f/2)$ Hz. The procedure was generally the same as that described for Experiment 1, except that there were only two possible ordered signal pairs. Trials were presented in blocks of 50 during which $|\Delta f|$ remained constant. The presentation levels and values of Δf were the same as those in Experiment 1, and the signals were again presented in a quiet background to the observer's left ear. Each test session, about one hour in duration, consisted of two series of eight blocks of trials. Half of these series of blocks were given with $|\Delta f|$ increasing over blocks, and half with that value decreasing. The presentation order of sessions at different presentation levels and with $|\Delta f|$ increasing and decreasing over blocks was counterbalanced within each observer. The observers each received at least 50 trials at each combination of signal level and $|\Delta f|$. No special training was given for this task, but each observer completed at least one hour of practice before the data presented below were collected, and had already participated in Experiment 1. No trial-by-trial feedback was given, but the observers were informed of their performance when they next attended. Again, two stimulus tapes

were prepared, each with the stimuli recorded in a different pseudo-random order, and each observer received each tape equally often; the two orderings of the change of $|\Delta f|$ over blocks were obtained by playing the tapes in both forward and reverse directions.

Statistical assumptions for Experiment 1.

Before the results of these two experiments are discussed, a number of aspects of the 2AFC decision model are considered with respect to the first of these experiments.

If it were assumed that the variances of the internal representations of the residue pitches of the two signals having $n=4$ and $n=6$ were equal, and that each representation could be adequately described by a unimodal normal distribution, the relative DLs estimated from the observed psychometric functions by the method described in Chapter 3 could be interpreted as estimates of the relative standard deviations of those representations. For several reasons, these assumptions may not be wholly justified. One difficulty with these assumptions is that while performance levels above 50% in the experimental task constitute strong evidence that the observers were responding on the basis of a difference in the perceived residue pitches of the two signals, it is not necessarily the case that these pitches were always in the region of the fundamental frequency of the signals, so that the assumption of unimodality is not necessarily correct. It can however be argued that if performance with a sufficiently large $|\Delta f|$ reaches 100% correct,

it is highly likely that at least the two pitches compared on a given trial were from the same mode of the distribution of the pitches of each signal, and thus rather unlikely that different modes were sampled on different trials. Such performance levels were attained by all three observers, and given the pre-test training received by the observers, the assumption of effective unimodality is plausible. The experimental task affords no means of testing the assumption that the variances of the internal representations of the pitch of the two signals are equal. Since the frequencies of the components of each possible pair of signals were different in frequency, this assumption may well be false, and its application would then lead to an estimate of $\sigma(f'_0)/f_0$ which was close to the larger of the two values of this parameter.

While this experimental task is not ideal for the estimation of $\sigma(f'_0)/f_0$, the experimental materials were primarily selected to provide a strong test of the perceptual reality of the residue pitch of a two-component harmonic complex for the available observers. Experiments 5 to 13, described in Chapters 5 to 8, have been performed with stimulus materials which are better suited to the estimation of $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$, and incorporate technical refinements whose value can be better appreciated through an understanding of the difficulties of interpretation met with here.

Results - Experiments 1 and 2.

Psychometric functions.

Figure 4.3 and 4.4 respectively show the observed psychometric functions from Experiments 1 and 2, where each function is the proportion of responses that the second signal was higher in pitch as a function of the signed value of $\Delta f/f_0$. The best-fitting normal ogives as given by a Probit analysis (Finney 1971) also appear in those figures. Since the fitted function is assumed to correspond to a cumulative distribution of differences between two distributions having equal variances, the relative DL is evaluated as the reciprocal of the slope of the fitted function divided by $\sqrt{2}$; this value will only correspond to a performance level of 76% correct if the observer employed a symmetric decision rule. The parameters of the best-fitting functions and the values of $\Delta f/f_0$ for each experiment appear in tables 4.5 and 4.6. The first table also gives details of the form in which the results of the subsequent Probit analyses will be tabulated. The deviation of the observed functions from the theoretical functions is generally greater than that expected from binomial variability, but since no systematic differences are evident, it is assumed that these deviations result from additional noise rather than from a basic inadequacy of the decision model.

Figure 4.3 Psychometric functions for Experiment 1. The figure shows the proportion of responses that the second signal was higher in pitch than the first, as a function of the signal value of $\Delta f/f$. The functions for different signal levels are displaced laterally, and the appropriate frequency-difference scale for each level is indicated to the right of the x axis. The solid lines are drawn through the best-fitting theoretical function derived from a Probit analysis.

Panel a) - observer A.F., panel b) - observer G.S., panel c) - observer B.P.

Symbol	S.L.	Trials / point		
		A.F.	G.S.	B.P.
Δ	15	200	100	100
O	20	200	100	100
■	30	200	100	100
G	40	200	100	100
Δ	50	200	200	100

Fig. 4.3 a) 0. A.E

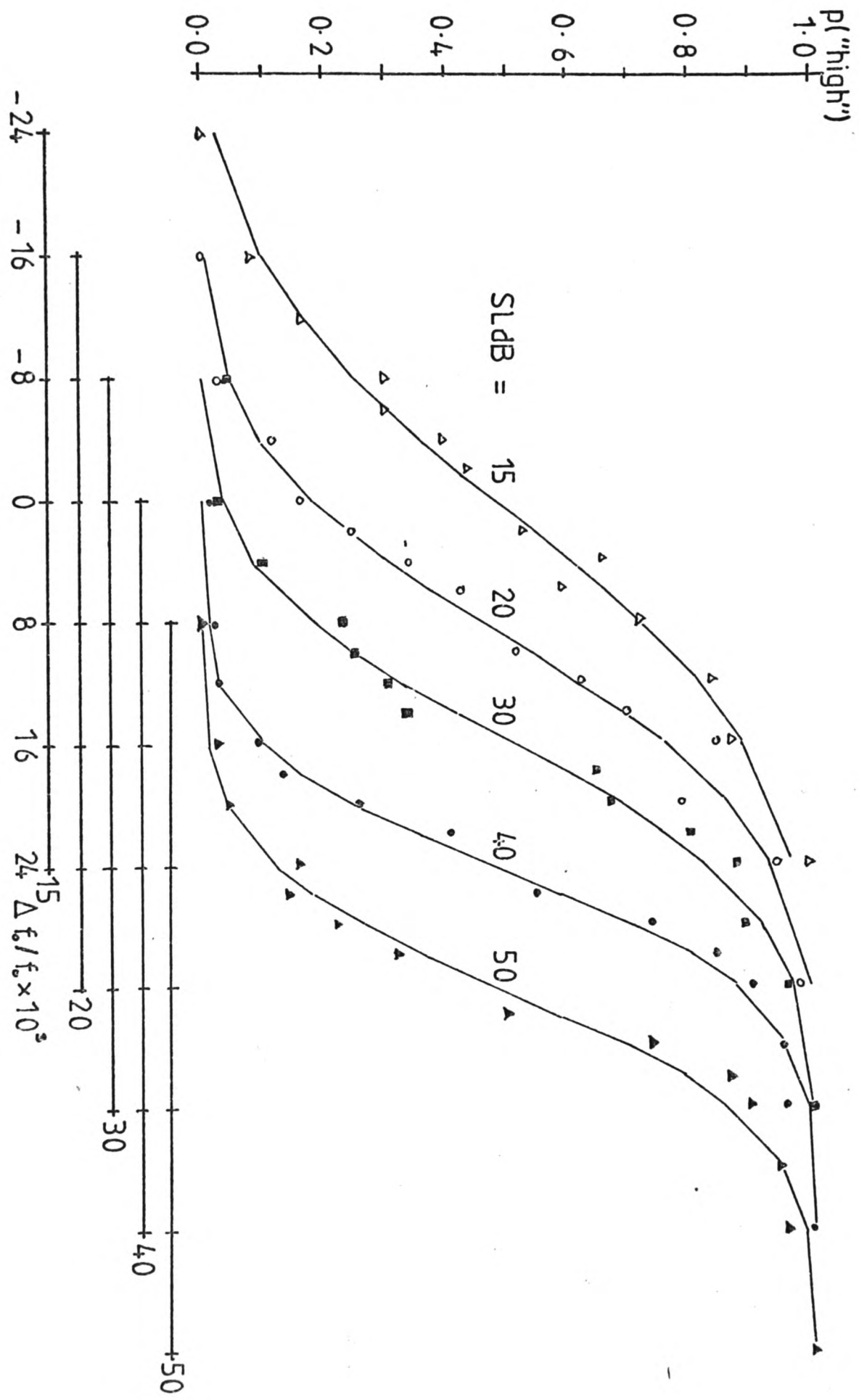


Fig. 4.3 b) 0. G.S.

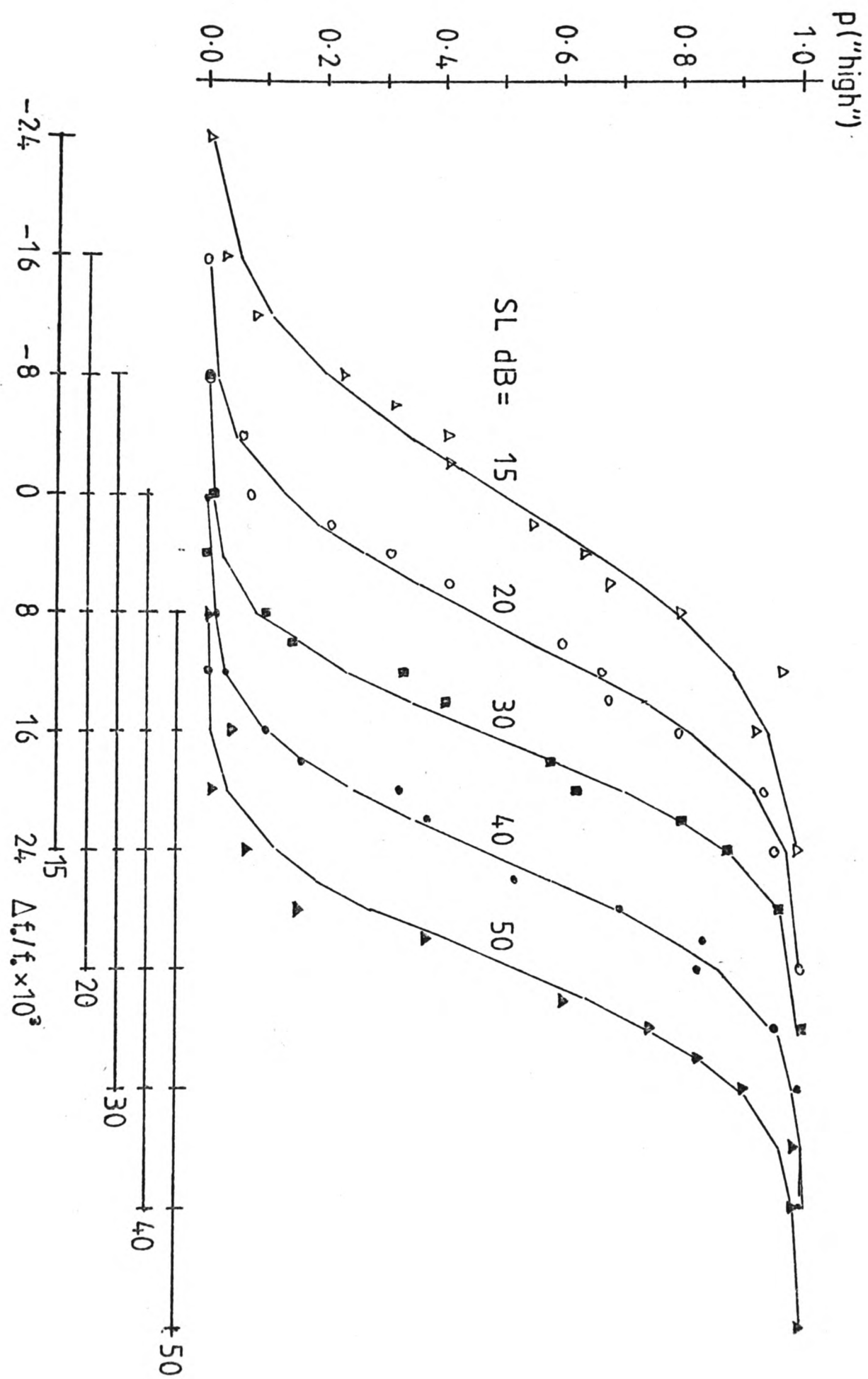


Fig. 4.3c) 0. B.P.

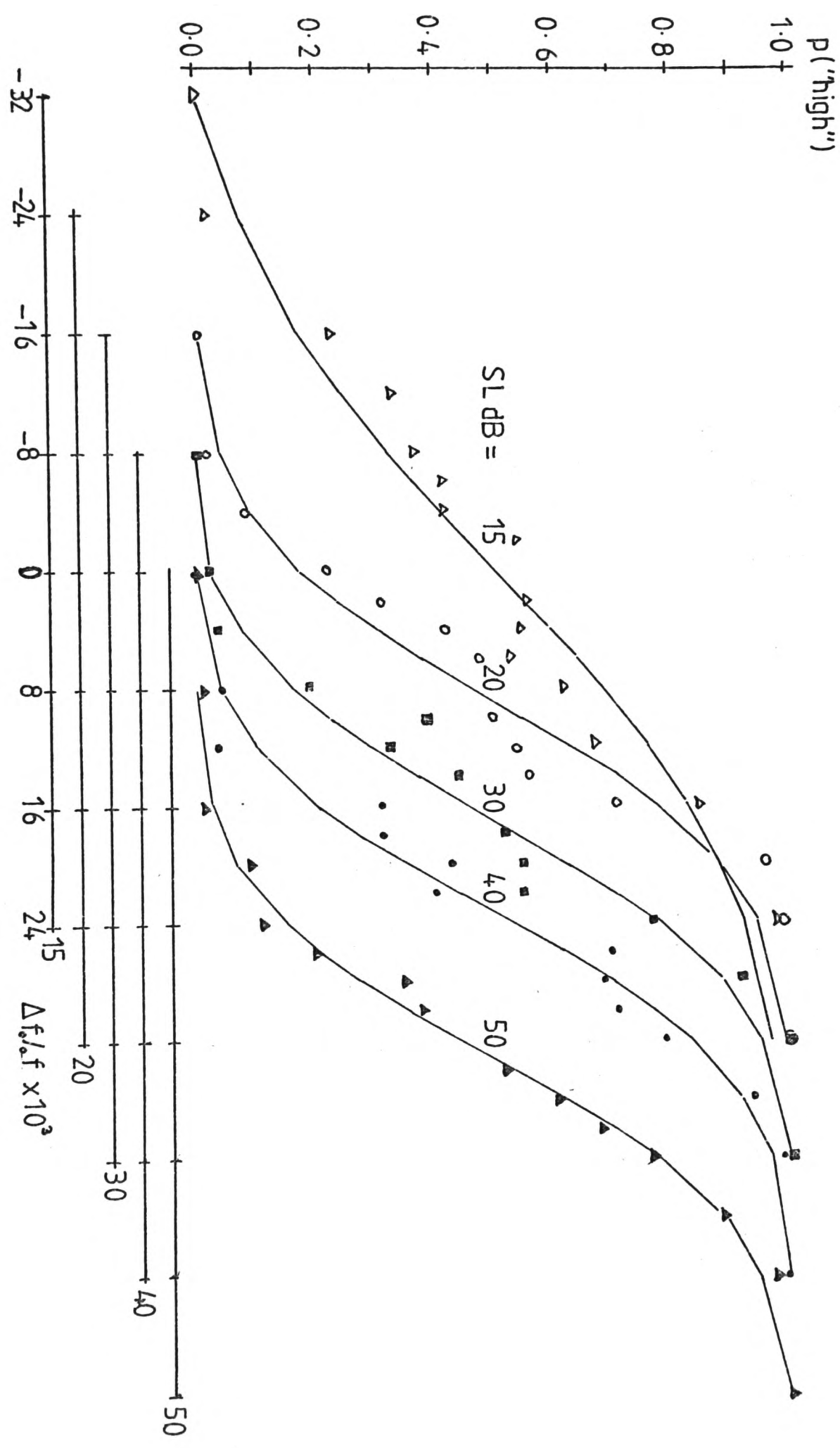


Figure 4.4 Psychometric functions for Experiment 2. The functions are laterally displaced as in figure 4.4. Panel a) - observer A.F., panel b) - observer G.S., panel c) - observer B.P.

Symbols	S.L.	Trials / point		
		A.F	G.S	B.P.
△	15	100	75	50
○	20	100	75	50
■	30	75	75	50
●	40	75	75	75
▲	50	75	75	50

Fig. 4.4a) 0. A.F.

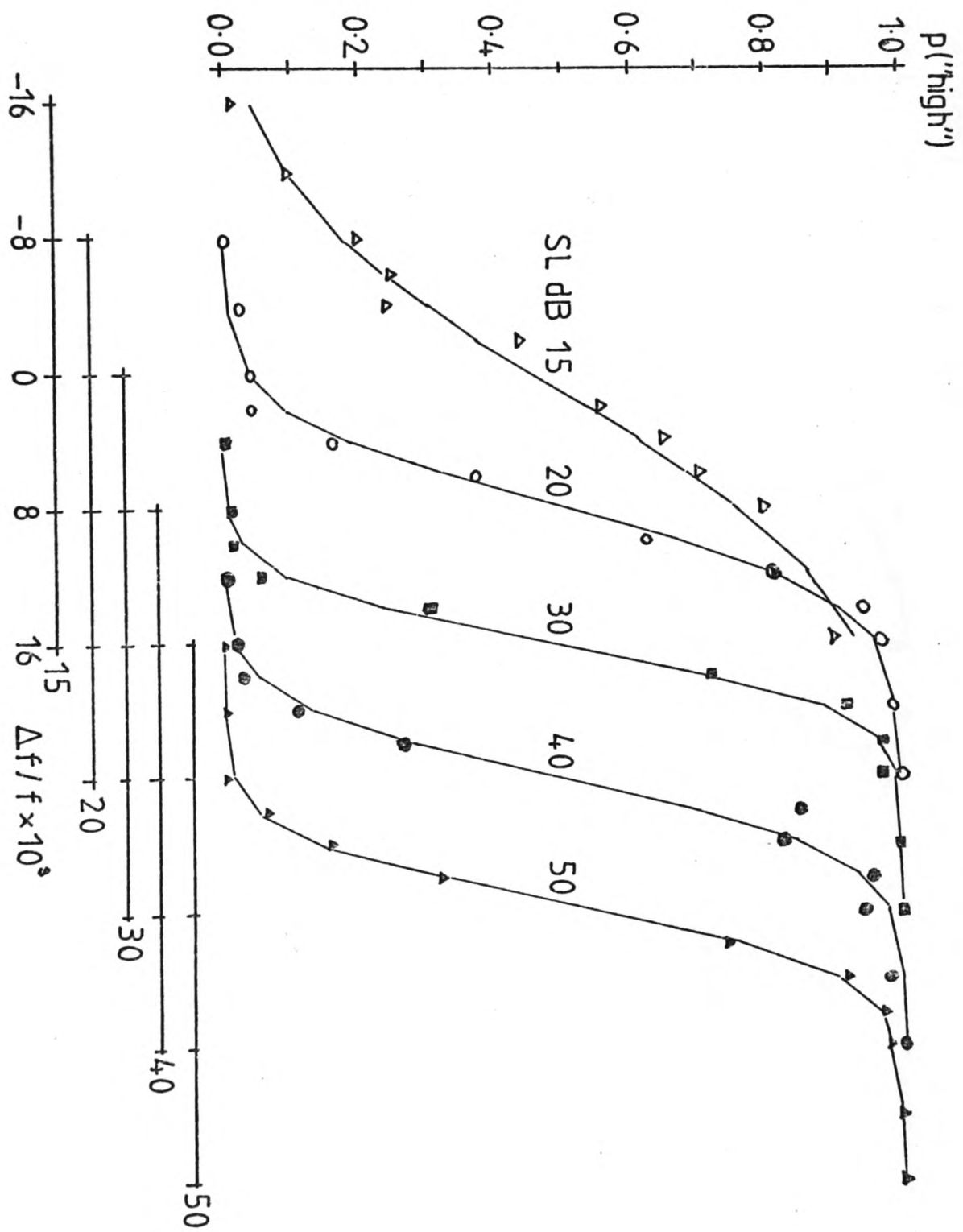


Fig. 4.4 b) 0.G.S.
P("high")

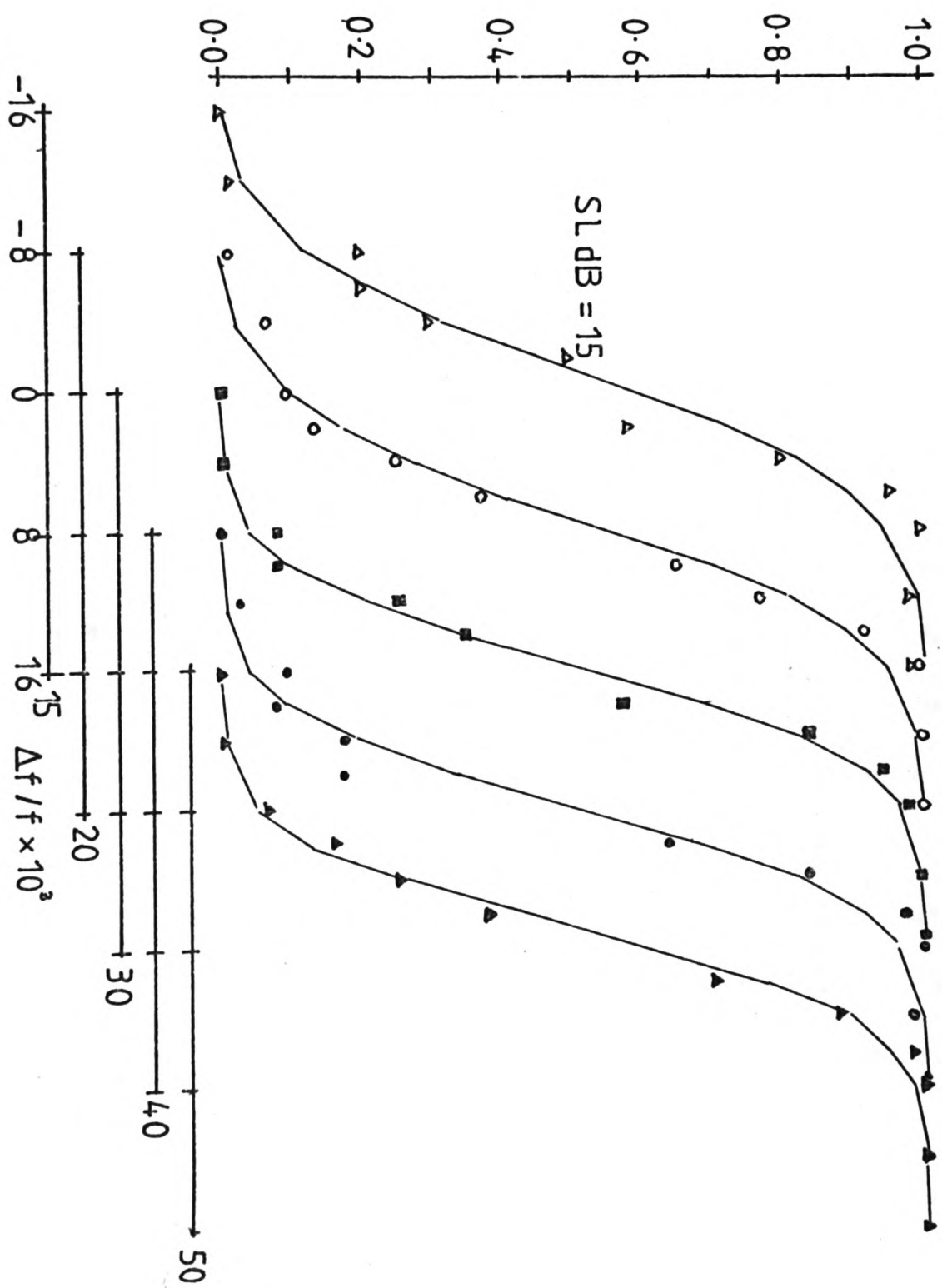


Fig. 4.4 c) O. B. P.

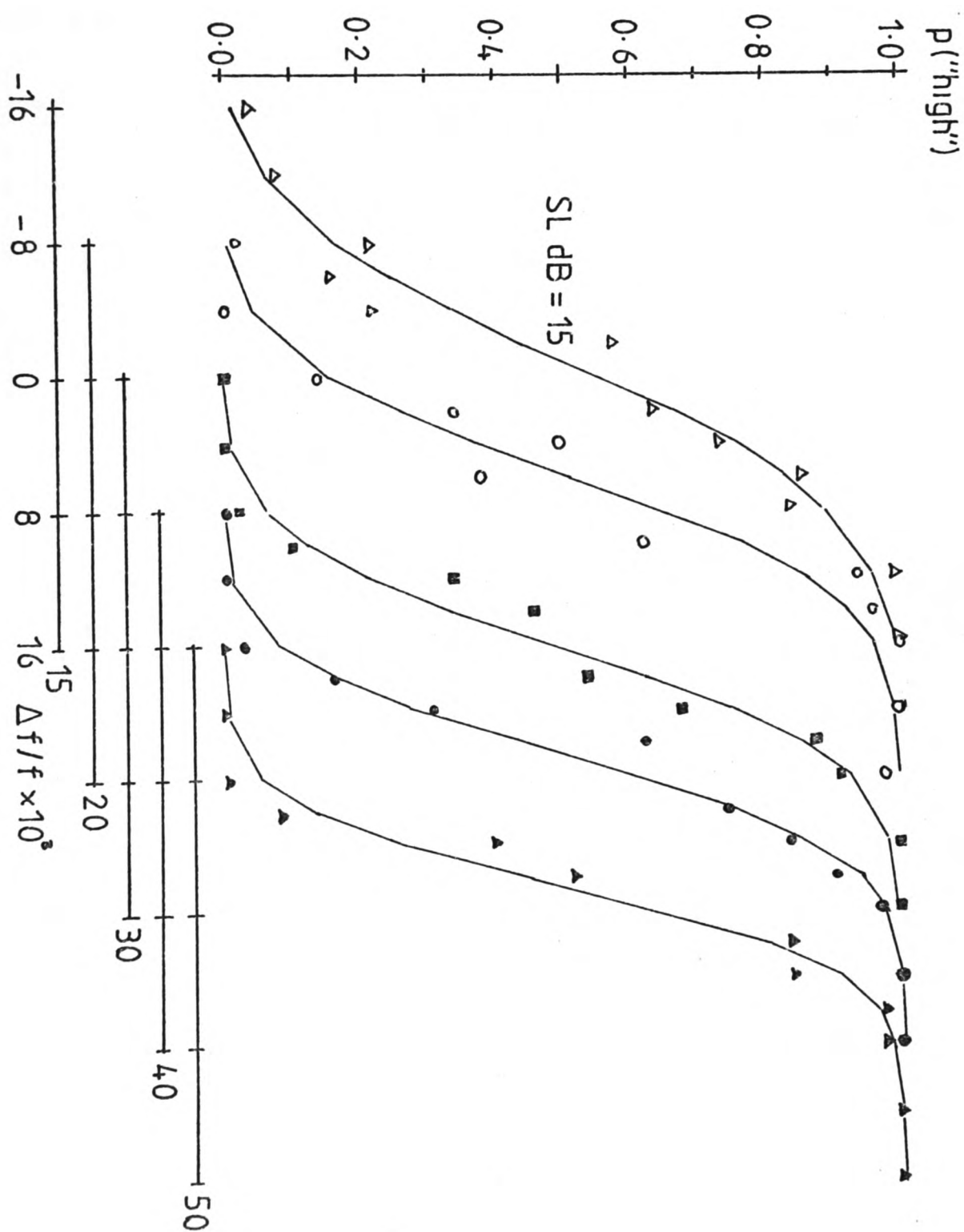


Table 4.5. Probit analysis of observed psychometric functions from Experiment 1. In this and all similar tables below, the column legends have the following meaning. N is the number of trials contributing to each point of the function for each signed value of $\Delta f / f_0$. The slope and 50% point are in units of $\Delta f / f_0$. The 50% point is that value of $\Delta f / f_0$ at which the theoretical function has a value of 0.5. S.E. is the standard error of the estimated slope and 50% point. df is the number of degrees of freedom for χ^2 , and hf the heterogeneity factor. The heterogeneity factor has a value of 1 if χ^2 does not exceed a $p=0.05$ criterion, and otherwise is χ^2/df . Values of χ^2 marked with an asterisk were obtained by combining observations to give a minimum of 5 or a maximum of N-5 "high" responses. Probabilities greater than 0.05 are not shown.

Observer	SL	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
A.F.	50	200	136.82	9.17	+0.0003	0.0005	30.919*	8	3.885	<0.01
	40	200	136.99	7.65	+0.0002	0.0004	21.252*	8	2.657	<0.01
	30	200	109.46	3.88	-0.0004	0.0003	12.307*	12	1.0	
	20	200	99.94	5.49	+0.0008	0.0005	27.755	12	2.313	<0.01
	15	200	77.58	3.145	+0.0004	0.0004	15.999	12	1.0	
G.S.	50	100	153.26	7.43	-0.0003	0.0003	7.895*	6	1.0	
	40	100	147.36	7.12	+0.0006	0.0003	9.240*	7	1.0	
	30	100	154.41	7.49	+0.0005	0.0003	13.633	14	1.0	
	20	100	125.84	6.03	+0.0008	0.0004	11.812	14	1.0	
	15	100	100.70	4.78	-0.0001	0.0004	14.459*	10	1.0	
B.P.	50	100	110.57	5.31	+0.0012	0.0004	10.728	14	1.0	
	40	100	108.71	6.83	-0.0009	0.0005	23.991	14	1.714	<0.05
	30	100	108.01	9.08	+0.0009	0.0007	43.649	14	3.118	<0.01
	20	100	103.79	8.69	+0.0005	0.0007	43.498	13	3.107	<0.01
	15	100	51.54	2.63	-0.0004	0.0007	18.631*	12	1.0	

Table 4.6. Probit analysis of observed psychometric functions from Experiment 2.

Obs.	SL	N	Slope	S.E.	50% point	S.E.	χ^2	df	hf	p
A.F.	50	75	288.76	18.00	-0.0005	0.0003	2.344	10	1.0	
	40	75	263.48	28.94	+0.0001	0.0005	32.627	10	3.263	<0.01
	30	75	310.86	19.74	+0.0001	0.0003	9.887	10	1.0	
	20	100	218.69	12.84	+0.0000	0.0003	17.633	10	1.0	
	15	100	88.84	6.04	+0.0010	0.0005	8.492	10	1.0	
G.S.	50	75	235.42	14.14	-0.0012	0.0003	6.927	12	1.0	
	40	75	220.82	22.04	-0.0001	0.0005	34.577	12	2.881	<0.01
	30	75	218.52	12.84	-0.0003	0.0003	11.400	12	1.0	
	20	75	180.61	10.32	-0.0009	0.0003	19.683	13	1.0	
	15	75	171.61	14.56	-0.0014	0.0005	28.874	13	2.210	<0.01
B.P.	50	50	245.32	18.37	-0.0015	0.0003	14.435	10	1.0	
	40	75	208.43	12.19	-0.0014	0.0003	16.399	13	1.0	
	30	50	183.27	12.83	+0.0002	0.0004	14.601	10	1.0	
	20	50	172.38	24.20	-0.0023	0.0008	19.830*	5	1.0	
	15	50	135.95	9.41	-0.0011	0.0005	15.387	10	1.0	

One particular difficulty in the interpretation of the results of this experiment becomes evident from a consideration of the proportions of the two possible responses made to each of the four possible signal pairs. In figure 4.5, separate psychometric functions are plotted for the signal pairs where the lower harmonic number of the first signal is 4 and for those pairs where the lower harmonic number of the first signal is 6. It is apparent that, for observers G.S. and B.P., there is a considerable difference in the intercepts of the psychometric functions for these two classes of signal pair. This intercept difference may well be partly responsible for the deviation of some of the overall functions from the cumulative normal form, since the sum of two normal distributions which overlap but have different means is not a unimodal normal distribution.

The effect of presentation level on discrimination for residue pitch.

Figure 4.6 shows for each observer the effect of presentation level upon the estimated standard deviation $\sigma(f'_0)/f_0$, where $\sigma(f'_0)/f_0$ is estimated from the psychometric functions from Experiment 1 pooled over all signal pairs. The error bars shown on this figure are 95% confidence limits, calculated as described by Finney (1971), and based upon the 95% confidence limits for the estimated slope of the psychometric functions. The confidence limits were calculated as ± 1.96 standard errors for the estimated slope if χ^2 did not exceed the $p=0.05$ criterion, and otherwise as $\pm t(v, 0.05)$ standard errors where v is the number of degrees of freedom for the evaluation

Figure 4.5 Psychometric functions from Experiment 1, with the data broken down into the two possible orderings of the signals. Empty symbols, dashed line - signal having higher harmonics presented in the first interval. Filled symbols, solid line - signal having higher harmonics presented in the second interval.

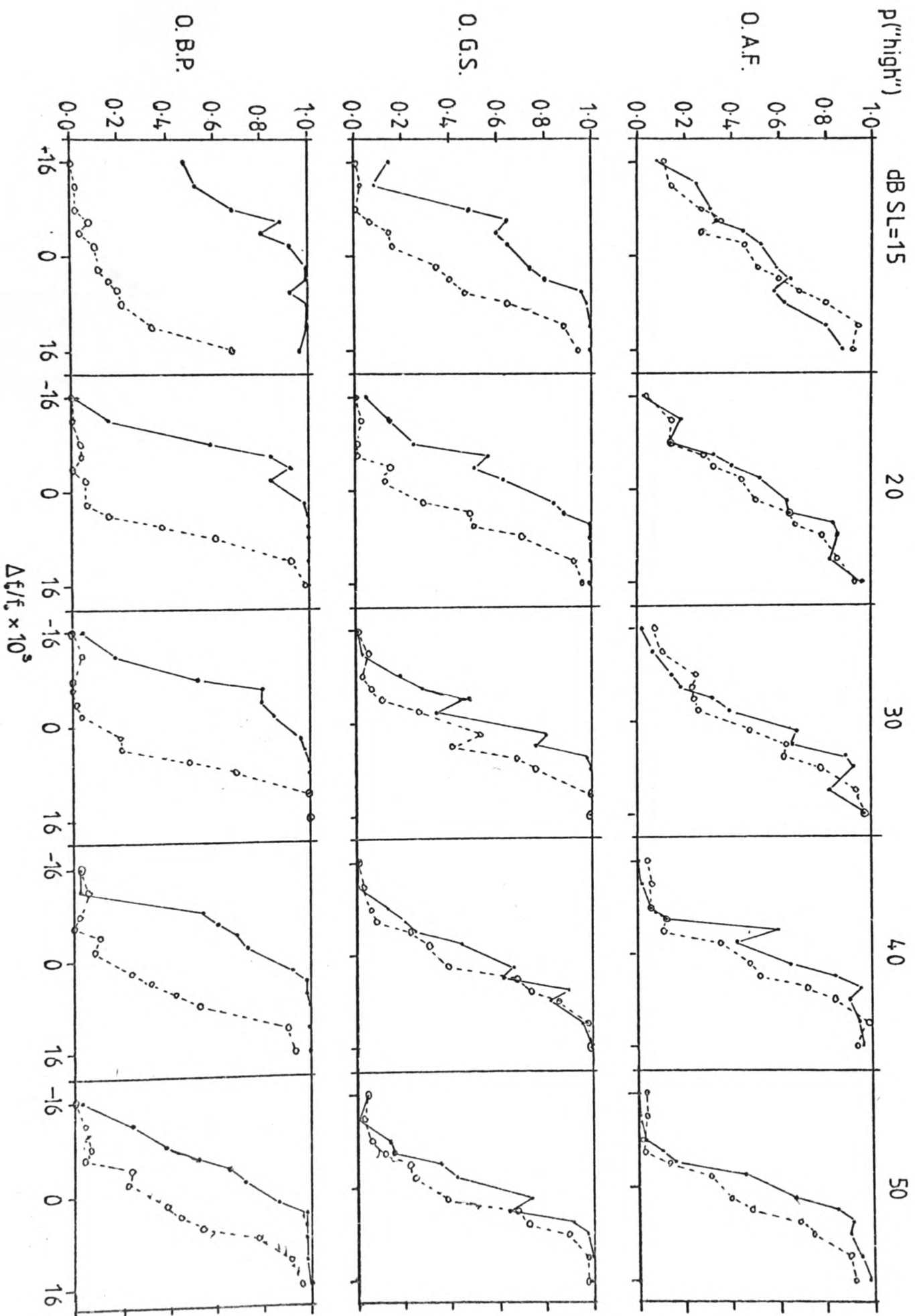


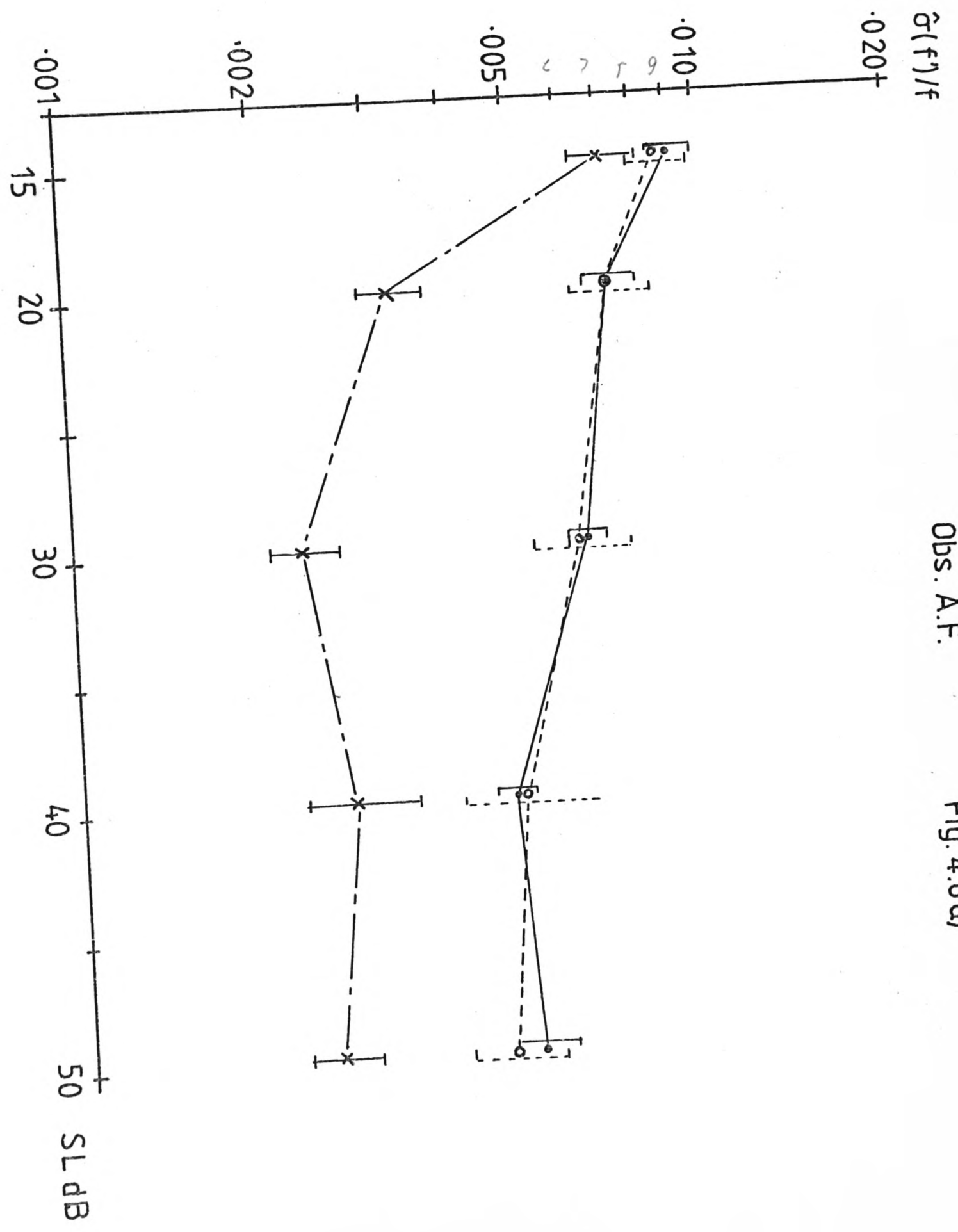
Figure 4.6 Estimates of $\hat{\sigma}(f'_0)/f_0$ and the relative frequency DL at 1256 Hz as a function of Sensation Level. Results from three observers, panel a) - observer A.F., panel b) - observer G.S., panel c) - observer B.P.

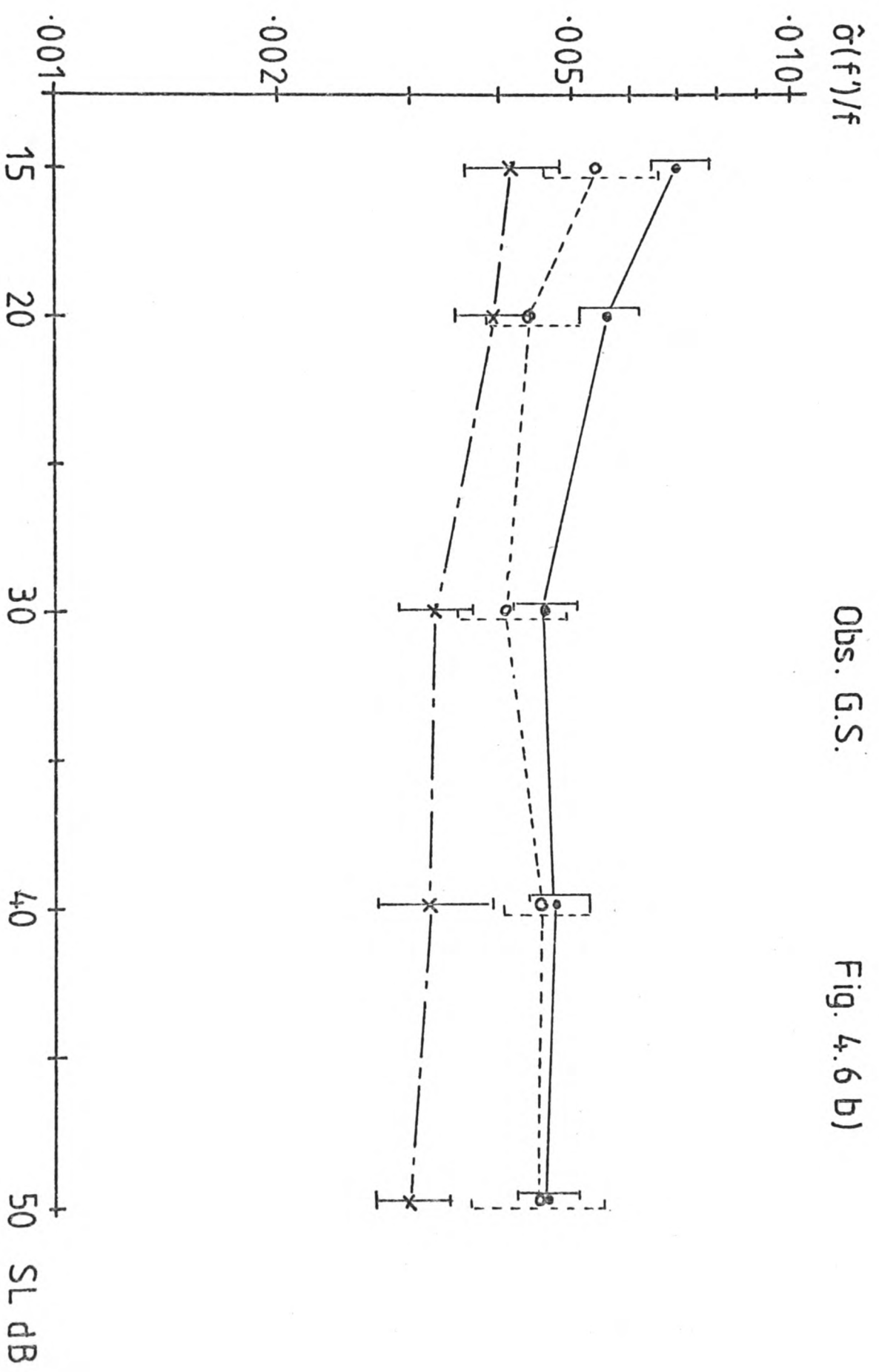
- $\hat{\sigma}(f'_0)/f_0$ pooled data.
- $\hat{\sigma}(f'_0)/f_0$ estimated from two stimulus orders separately.
- × $\Delta f / f$ at 1256 Hz, from Experiment 2.

The error bars indicate 95% confidence limits.

Obs. A.F.

Fig. 4.6a)

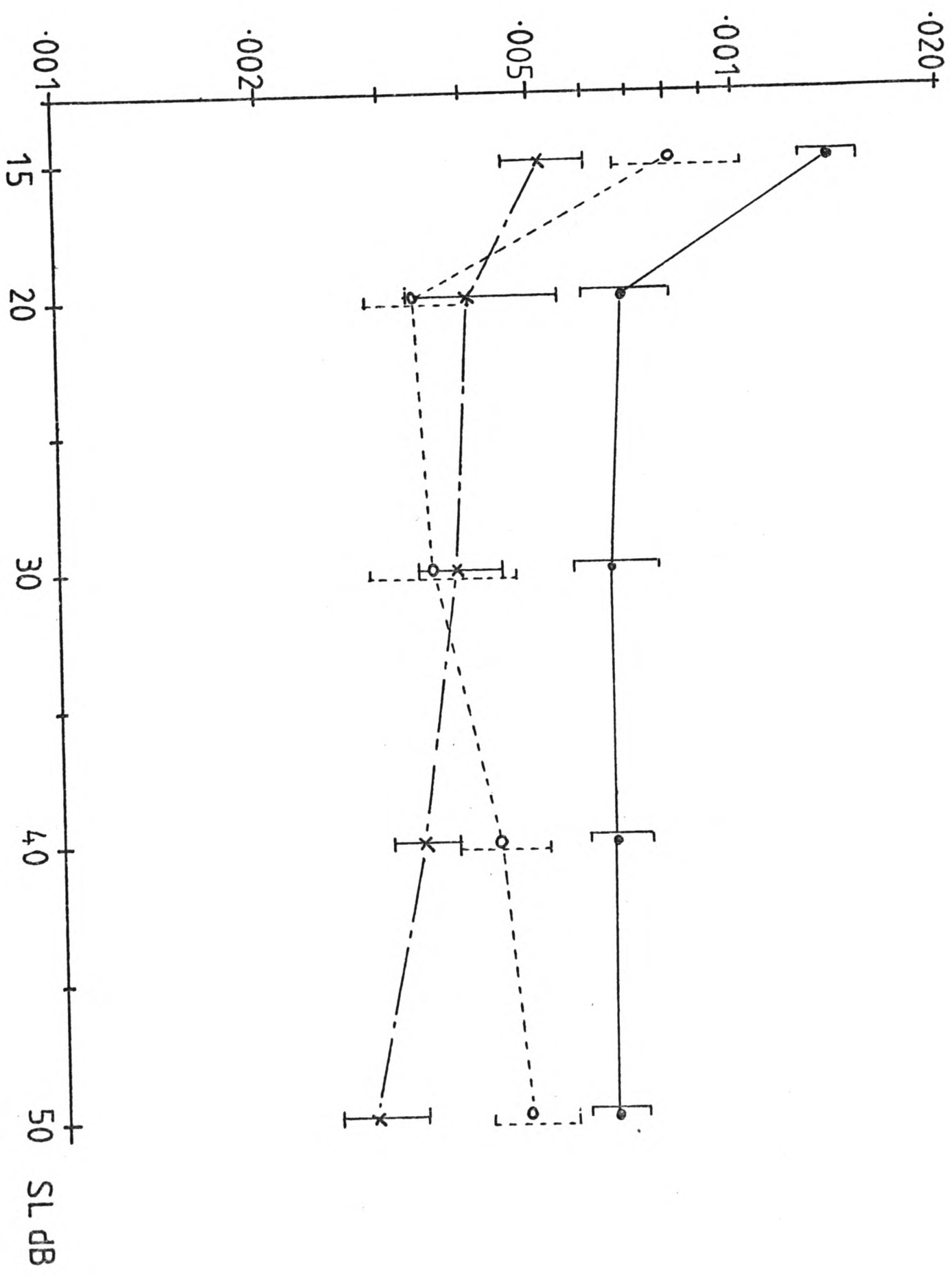




$\hat{\sigma}(f')/f$

Obs. B.P.

Fig. 4.6 c)



of χ^2 . If χ^2 exceeds the $p=0.05$ criterion, the Probit analysis evaluates the variance of the estimated slope as the product of the variance due to the binomial distribution and a heterogeneity factor, where the latter term is given by χ^2/df . All subsequent graphs of estimated relative standard deviations show similarly calculated confidence limits.

Figure 4.6 also shows estimates of $\sigma(f'_0)/f_0$ derived from separate functions for the two orderings of the signals of different harmonic number. These estimates tend to be smaller than those derived from the pooled data, particularly where the differences in the intercepts of the functions for the two signal pairs are greater; the confidence limits of these estimates, however, are rather larger.

As would be expected from the spectral-pattern recognition models, the level dependence of the relative DL for the residue pitch of the complex signals presented in Experiment 1 is roughly comparable with that for the single sinusoid presented in Experiment 2. The effects of presentation level are comparable with those found by Weir, Jesteadt, and Green (1977) for single sinusoids, and, like those data, follow approximately the form of the amplitude dependence predicted by both Siebert's (1970) temporal model and by the modified Energy Detector models described in Chapter 3.

The first of these experiments does suggest that $\sigma(f'_0)/f_0$, and hence $\sigma(f'_j)/f_j$, contrary to Goldstein's assumption, are both dependent upon factors other than

frequency, and show a similar amplitude dependence to that shown by the relative frequency DL for single sinusoids. The values of the estimates of $\sigma(f'_0)/f_0$ are somewhat smaller than would be appropriate to the estimate of $\sigma(f'_j)/f_j$ derived by Goldstein (1973) in fitting his model of pitch perception to the error rates in the musical interval identification data of Houtsma and Goldstein (1972), where a value of 0.01 was a suitable estimate of $\sigma(f'_j)/f_j$. Houtsma (1979), from a different body of musical interval identification data, estimated from the error rate a value for $\sigma(f'_j)/f_j$ of about 0.011. These values of $\sigma(f'_j)/f_j$ should lead to a minimum relative DL for the residue pitch of comparable two-component complexes of about 0.008, whereas the smaller relative DLs for Experiment 1 are approximately 0.005, corresponding to a value for $\sigma(f'_j)/f_j$ of about 0.007. From the simulations of Goldstein's model described in Chapter 2, the error rate in assigning a residue pitch to the complex having $n=4$ should be about 0.02, so the error rate in identifying a musical interval represented by the fundamental frequencies of two such signals should be about $1-(0.98^2)$, or about 0.04. Houtsma and Goldstein (ibid) and Houtsma (ibid) reported error rates for comparable signals of between 0.05 and 0.10, which are approximately consistent with the estimates of $\sigma(f'_0)/f_0$ obtained here. One reason for the discrepancy between the present estimates of $\sigma(f'_0)/f_0$ and the estimates of $\sigma(f'_j)/f_j$ from musical interval identification may be that the latter task is subject to errors from sources other than pitch processing.

CHAPTER 5

Abstract.

This chapter describes an attempt to estimate experimentally the standard deviations of the internal representations of the frequencies of each of the components of a two-component complex signal. It is assumed that these internal representations are identical to the component frequency estimates which, in the spectral-pattern recognition models described above, determine the residue pitch pdf. The statistical methods employed were those described in detail in Chapter 3 for a 2AFC frequency discrimination task. Frequency discrimination as a function of the presentation level of the components of the complex was measured for each component separately, both components being presented together. The component whose frequency was to be discriminated was presented as two temporally successive tone-bursts, while the other component was continuously present. The simultaneous presentation of these two components was intended to allow an estimate of the relative frequency DL which included any variability due to the second component, and which would thus be a more appropriate estimate of $\sigma(f'_j)/f_j$ than that provided by the relative frequency DL for a single sinusoid.

Initially, the stimuli were presented in a silent background, so that presentation conditions would be similar to those for the residue pitch discrimination task described in the previous chapter, but it was found that this procedure led in some circumstances to psychometric functions which were almost flat. This difficulty was ascribed to the influence of odd-order auditory distortion products, and the experiment was repeated with the stimuli presented in a background of noise, low-pass filtered so as to mask any such distortion products. The observed relative difference limens were not a simple monotonically decreasing function of presentation level, but decreased with increasing level for signals up to 30 dB SL, and showed a marked increase at 50 dB SL. This non-monotonicity was attributed to two factors. For signal levels greater than 30 dB SL, the masking noise led to a constant signal-to-noise ratio, which would be expected to give rise to a constant DL. The increase in the DL at 50 dB SL was assumed to arise from an increased upward spread of masking from the low-pass filtered noise.

Introductory Comments.

Apart from the studies of the identifiability of the components of complex signals described by Plomp (1964), Plomp and Mimpen (1968) and Moore (1973b), which were mentioned in Chapter 1, there have been no psychophysical studies of the accuracy of a listener's internal representation of the frequencies of a complex signal. One testable property of the spectral-pattern recognition models of the perception of residue pitch is that, as was shown in Chapter 2, the standard deviation of the internal representation of that pitch is directly dependent upon the standard deviations of the internal representations of the components of the complex signal. The above-mentioned studies of the identifiability of the components of a complex signal yield estimates of the minimum frequency difference between adjacent components at which the standard deviation of the internal representation of a component is small with respect to this frequency difference. They do not, however, provide quantitative estimates of this standard deviation where the separation between the components of the signal is sufficient for the resolution of the components.

If the auditory system codes the frequencies of several well-resolved components of a complex signal independently, then frequency discrimination for each of these components should show the same dependence upon variables such as signal level, signal-to-noise ratio and signal duration as that shown for a single sinusoid. If, however, there is some interaction between the coding of the various components, these dependencies may

differ.

Problems of Experimental Method.

If it may be assumed that the theoretical parameter $\sigma(f'_j)/f_j$ is primarily determined by external noise and by the variability of peripheral encoding, then this parameter may be estimated as the relative frequency difference limen for the component at f_j in the presence of the other components of the complex signal. The design of a suitable task for this purpose poses certain technical problems. In order to ensure that the experimentally obtained estimate includes the effect of components other than the j th, this component must be presented together with the other components. If this is done, however, the observer's task involves not just a frequency discrimination judgement, but also an attentional one. The attentional task involved here becomes apparent from a consideration of the multiple-band Energy Detector model described in Chapter 3. If an observer is to estimate the frequency of some sinusoidal signal, it is to his advantage to exclude as much noise as is possible. If, as is generally accepted, the peripheral stages of the auditory system can be described in terms of a series of parallel band-pass filters, each of which can contribute noise to the decision process, then the observer may minimize the contributed noise by attending to only those filters which convey information about the signal.

One study which involved a similar problem has been described by Schodder and David (1960). Here, observers heard,

in two successive temporal intervals, two signals, each containing two components. One component was of the same frequency in both intervals, while the second differed in frequency between the two presentations; both components were tone pulses and were of equal duration. Schodder and David were interested in whether observers compared the pitches of the two signals on the basis of the change in the frequency of the variable component, or upon the change in the envelope rate of the complex, which changes in the opposing direction to the change of the frequency of the variable component if that component is the lower. The results of this study are not of theoretical interest here, but this method, with some modification, has been adopted for the experimental investigation of frequency discrimination for a single component of a two-component complex described here. One problem which arises in the interpretation of Schodder and David's study is that different observers appeared to have employed different aspects of the signals as the bases for their judgements, and no attempt was made to draw the attention of the observers to the change in the frequency of the variable component of the signal. The method described below overcomes this difficulty by means of the continuous presentation of the one component of constant frequency, and the addition to this component of the variable component as a tone pulse. Since the constant component is continuously present, the observer is better able to attend to differences between the two sinusoidal tone pulses. As is evident from the results of this experiment, where the confounding role of auditory distortion products is eliminated by a masking noise, the difficulties encountered with Schodder

and David's study are absent with this modification of their method.

The Adopted Experimental Method.

This experiment employed the 2AFC frequency discrimination task described in Chapter 3, where the observer's task was to judge the pitch of the second test stimulus as higher or lower than that of the first. The two test stimuli were presented against a background of a continuous 1047 Hz or 837.6 Hz sinusoid. The test stimuli were equal in duration and intensity, so that it can be reasonably assumed that $\sigma(x'_1) = \sigma(x'_2)$. In order to estimate the standard deviation of the representation of the frequency of the test signals, it is necessary to make some assumption about the covariance of the two representations. Since the variances of these two representations may be influenced by the background to the presentations, which includes a continuous periodic component, it is plausible that the covariance may not be zero; if the covariance is unknown but not zero, the standard deviation of the representation of frequency cannot be estimated. In order to allow the estimation of this parameter, it will be assumed that the covariance is negligible, although the possibility that the estimate will be in error cannot be neglected. If the covariance were incorrectly assumed to be zero when its true value was greater than this, the resulting estimates of the standard deviation of the internal representation of frequency would be too small.

Observers.

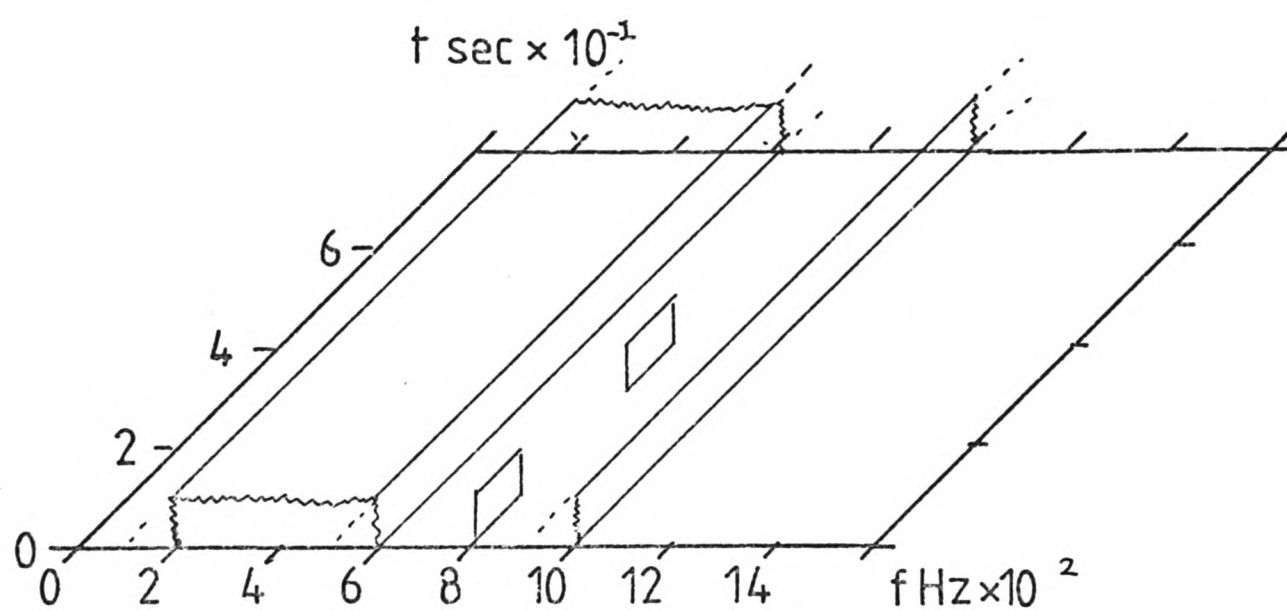
Two observers, G.S. and A.F., who had both participated in the experiments described in the previous chapter, also participated in this experiment.

Stimuli.

The two pulsed sinusoids presented during the two temporal intervals of each trial had the frequencies $\underline{n}(209.4 + \underline{\Delta f}/2)$ and $\underline{n}(209.4 - \underline{\Delta f}/2)$ Hz, where \underline{n} was equal to 5 during the first series of test sessions, and equal to 4 for the second series of sessions. The component at $\underline{m}(209.4)$ Hz, where $\underline{m}=4$ if $\underline{n}=5$ and $\underline{m}=5$ if $\underline{n}=4$, was present continuously throughout each test session. The absolute value of $\underline{\Delta f}$ ranged from $0.03248 \times \underline{n}(209.4)$ to $0.00101 \times \underline{n}(209.4)$ Hz. The time-course of a trial is illustrated in figure 5.1.

The intensities of the two pulsed sinusoids and the continuous sinusoid were always equal, and such that the 1047 Hz component was at approximately 15, 20, 30, 40, or 50 dB SL. Except where noted, these signals were presented with a continuous background of noise, which was low-pass filtered at 618 Hz with a filter slope of about 90 dB/octave, obtained by the use of four 24 dB/octave filters in series. The average noise spectral-density (that is, the average noise power in a band 1 Hz wide) in the region of 618 Hz was approximately 15 dB below the levels of the components at 837.6 and 1047 Hz. Because the level difference between the low-pass noise and the

Figure 5.1. The time-course of a single trial for Experiment 3. The block between the frequencies 200 and 600 Hz represents the low-pass noise.



signal needed to be constant to ensure the masking of any cubic-distortion product, and it was impossible to remove all of the noise energy in the region of the signal component frequencies, the signal-to-noise ratio in this frequency region was constant at the higher signal levels, while at the lower signal levels, the level of the noise in this frequency region was below the noise floor of the equipment. The measured signal and noise intensities at the two component frequencies and at 618 Hz at the signal levels used here are given in table 5.1.

Table 5.1. Signal level and noise spectral-density (N_0) in dB relative to a 0 dB SL sinusoid at 1047 Hz		
Signal level	N_0 between 837 and 1047 Hz	N_0 at 618 Hz
+50	+8	+35
+40	-1	+26
+30	-8	+12
+20	-10	+3
+15	-9	-5

Signal and noise levels were measured electrically at the headphone input with a Hewlett-Packard 3580A wave analyser and a bandwidth of 1 Hz. The equipment used for presenting the signals had a noise floor at about -9 dB relative to the reference level of 0 dB SL at 1047 Hz.

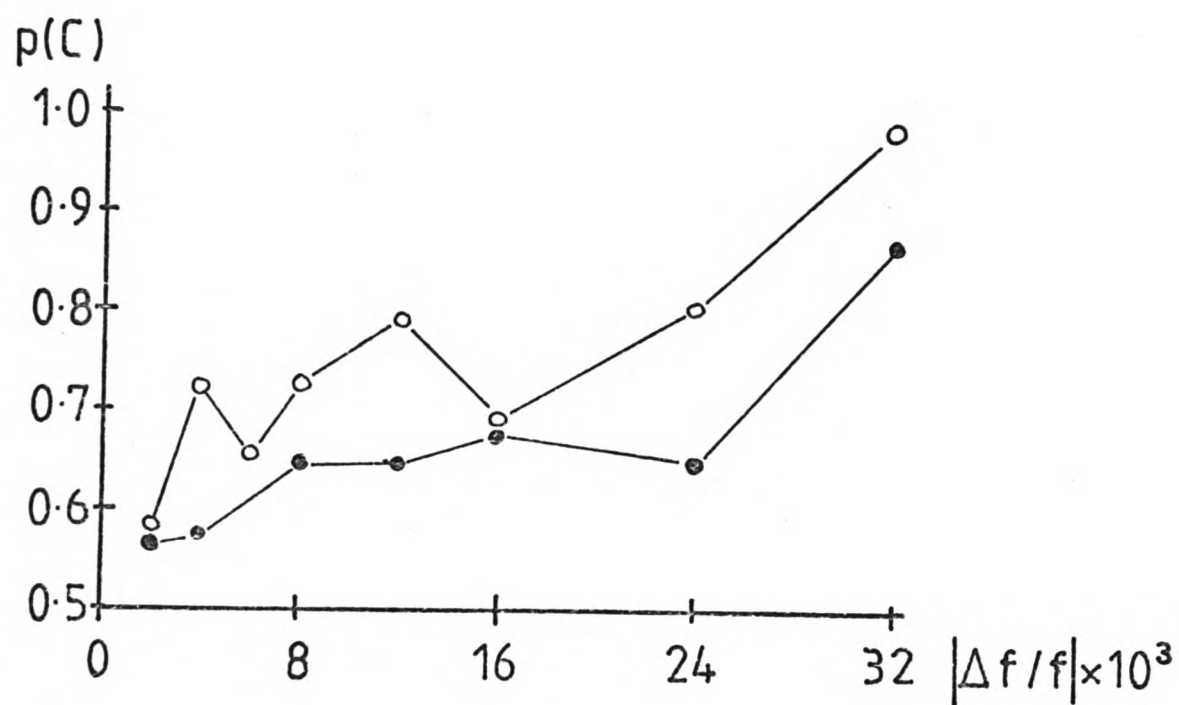
Procedure.

Trials were presented in blocks of 50. The absolute value of Δf was constant within each block, and the probability that the second pulsed sinusoid was higher in frequency than the first was 0.50; the order of trials within each block was determined pseudo-randomly. Sets of eight blocks at a constant SL were given with $|\Delta f|$ equally often increasing and decreasing over blocks. A set of blocks lasted about 25 minutes, and two sets of blocks at different SLs and with a different direction of change of $|\Delta f|$ were given in each test session. At least four sets of blocks at the same SL and with the same test

signals were presented to each observer, and the order in which these sets of blocks were given was approximately counterbalanced over the days of testing so that any day-to-day variability would be distributed over the different conditions. Two stimulus tapes were prepared, with the stimulus pairs in different pseudo-random orders. Reversal of the direction of the change in $|\Delta f|$ over blocks was achieved by reversing the direction in which the tapes were replayed.

Initially, the experiment was run without the background noise. Here, in the case of observer G.S. and for the test signal at 1047 Hz, the resulting psychometric functions at the higher SLs were approximately flat for $|\Delta f|$ less than $0.02436 \times n(209.4)$ Hz. Two examples of these functions are shown in figure 5.2. This lack of a dependence of performance upon Δf was not shown in this observer's performance with test signals at 837.6 Hz, nor in the data collected from observer A.F. This result might be explained with the assumption that an auditory nonlinearity generates a distortion product at the frequency $2f_2 - f_1$, where $f_1 = 837.6$ Hz and $f_2 = 1047 + 5(\Delta f/2)$ Hz. The level dependence of this distortion product, the cubic-difference tone, (Smootenburg, 1972a) is such that it may be expected to be audible at the higher SLs employed here. The frequency of this distortion product will be $618.2 - 5(\Delta f/2)$ Hz, so that this component changes in frequency in the opposite direction to the change at 1047 Hz. If observer G.S. was basing his judgements on frequency changes in both of these frequency regions, he would be using two conflicting cues, so that his performance could indeed be independent of Δf . If such

Figure 5.2. Two of the psychometric functions from Observer G.S. for Experiment 3 without the low-pass masking noise. The test signal is at 5×209.4 Hz. The plotted data are the probability of a correct response, $p(c)$, as a function of the absolute relative frequency difference. Filled symbols, test signal at 40 dB SL, open symbols, 30 dB SL.



a distortion product is generated by the signals used here, it will also give rise to difficulties where the test signal is at 837.6 Hz; in this case, the distortion product will change in frequency by $2 \times 4(\Delta f/2)$, which is twice the change at 837.6 Hz, and here the use of the distortion product as a cue would be expected to lead to an inflated performance score.

The low-pass filtered noise background was introduced to mask this distortion product. The spectral-density of the masking noise at 618.2 Hz was set at 15 dB below the levels of the components at 837.6 and 1047 Hz, which should completely mask the distortion product. At the frequencies of the two signal components, the noise spectral-density was either at least 40 dB below the level of the components or at a level at or below 0 dB SL (see table 5.1). Informal observations by observer A.F. confirmed that this level of noise was sufficient to render inaudible a 618 Hz sinusoid at a level 15 dB below the level of a 50 dB SL tone at 1047 Hz (the latter tone was not present during these observations).

After the experiment had been completed with the noise background, it was found that at SLs of between 20 and 40 dB, performance at the smallest values of Δf was approximately 70% correct, while at SLs of 15 and 50 dB, performance was close to 50% correct. In order to obtain more complete psychometric functions at the intermediate SLs, the experiment was repeated at 20, 30 and 40 dB SL with a minimum for $\{\Delta f\}$ of $0.00101 \times n(209.4)$ Hz. The availability of two sets of psychometric functions at these levels also permits an

assessment of the degree of improvement in the performance of the two observers with practice.

Results.

Psychometric functions.

The observed psychometric functions for the conditions with background noise were subjected to a Probit analysis. The functions fitted were those for the proportion of responses that the second test signal was higher in pitch as a function of the signed value of $\Delta f/f$.

The observed and best-fitting theoretical psychometric functions appear as figure 5.3; the parameters of the theoretical functions appear in table 5.2. The data from the repetition of the experiment with a minimum $|\Delta f|$ of $0.00101 \times f$ are indicated in both the figures and the table by the legend "replication". Of the 16 functions obtained for each observer, there are only three cases where the observed function is significantly different from a cumulative normal ogive with binomial variance with a criterion of $p=0.05$. All of these instances are functions obtained from observer G.S., and occur for $n=4$ at 40 dB SL (first set), 20 dB SL (replication), and at 15 dB SL. In all three cases, the observed function deviates from the theoretical function in that at larger absolute values of Δf , there are slight non-monotonicities, and such deviations can be attributed to some source of variability which is not taken into account by the decision model. Hence,

Figure 5.3. Psychometric functions from Experiment 3.

Panel a) observer A.F., $\underline{n} = 4$.

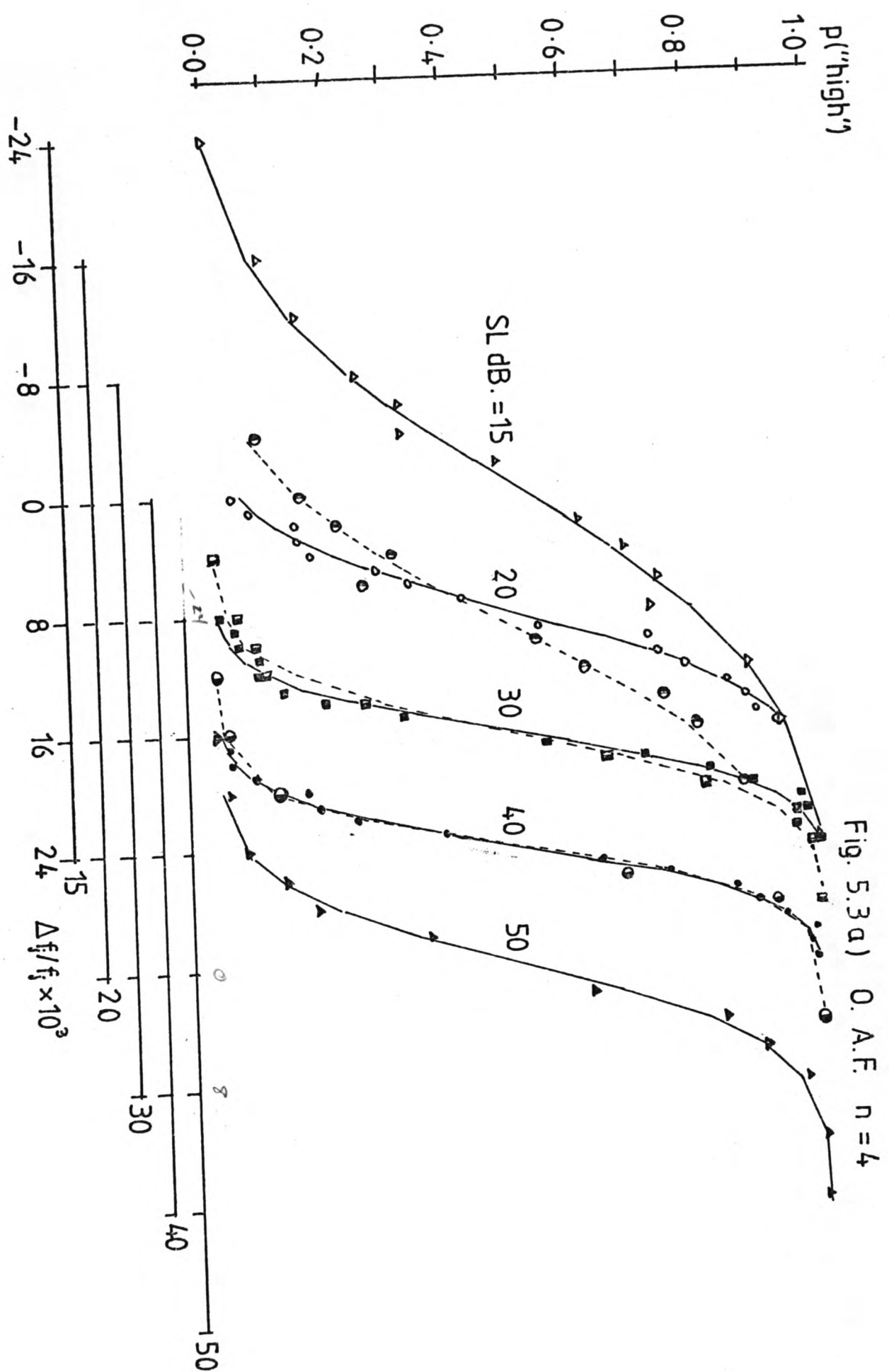
Panel b) observer A. F., $\underline{n} = 5$.

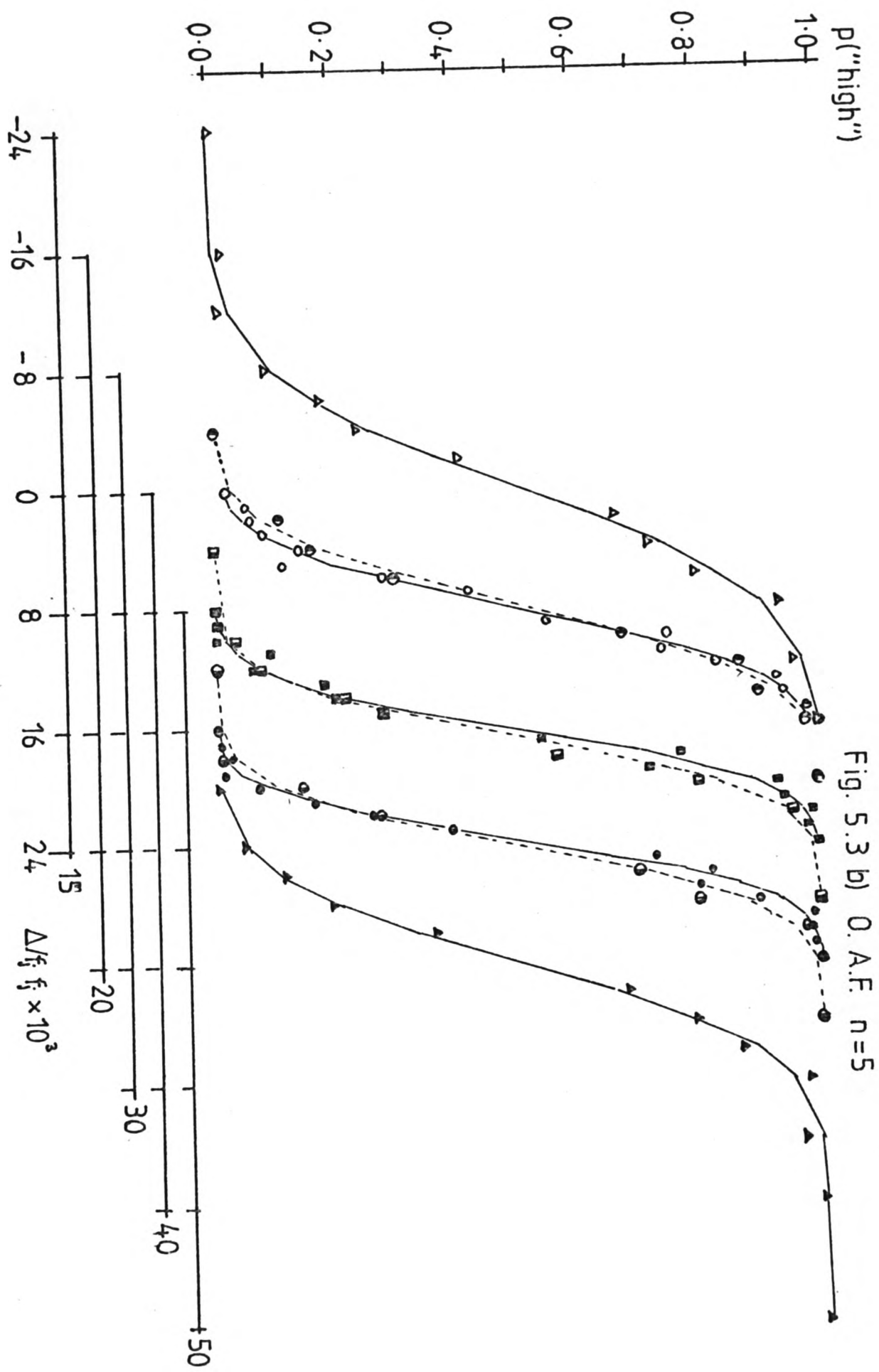
Panel c) observer G.S., $\underline{n} = 4$.

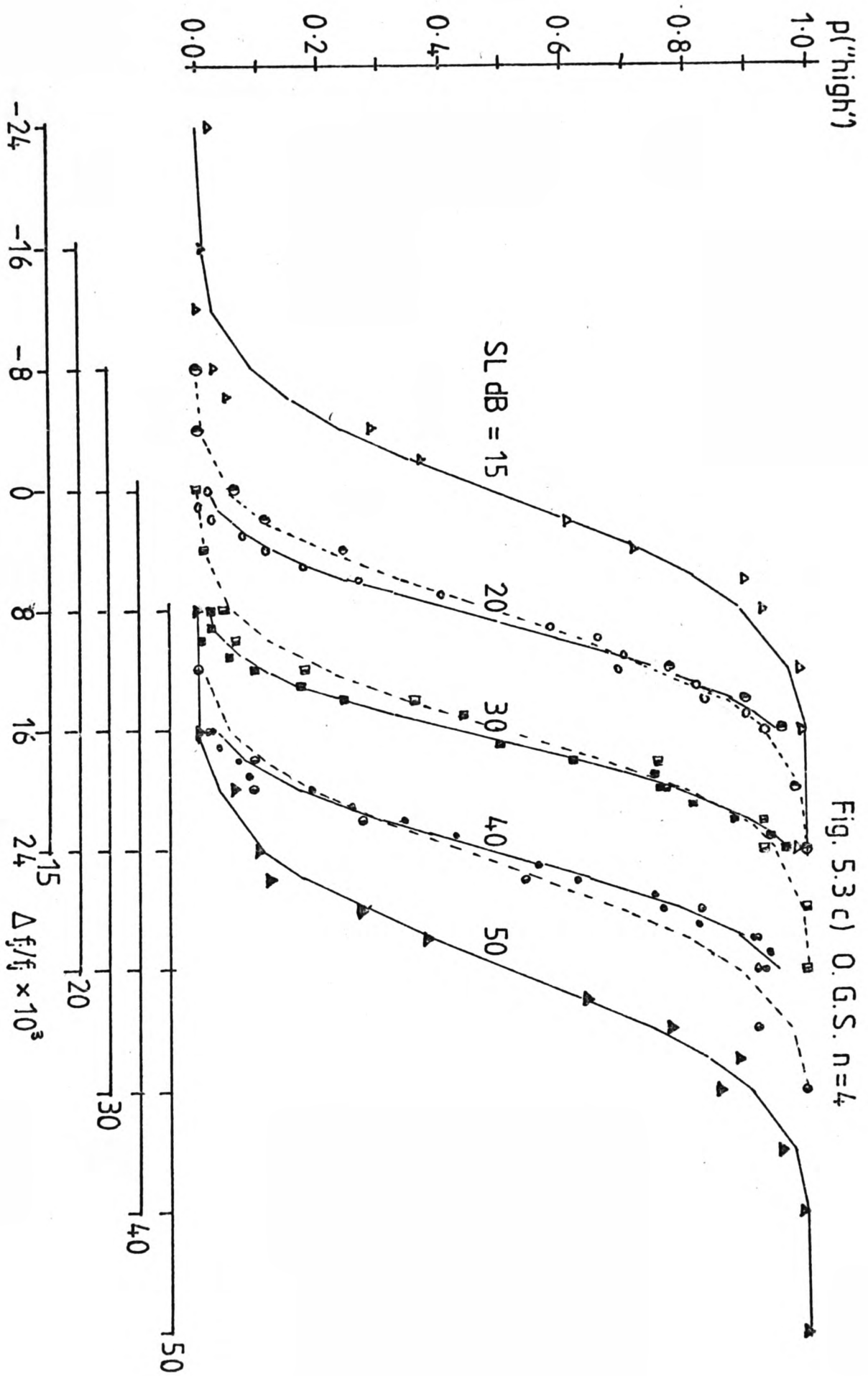
Panel d) observer G.S., $\underline{n} = 5$.

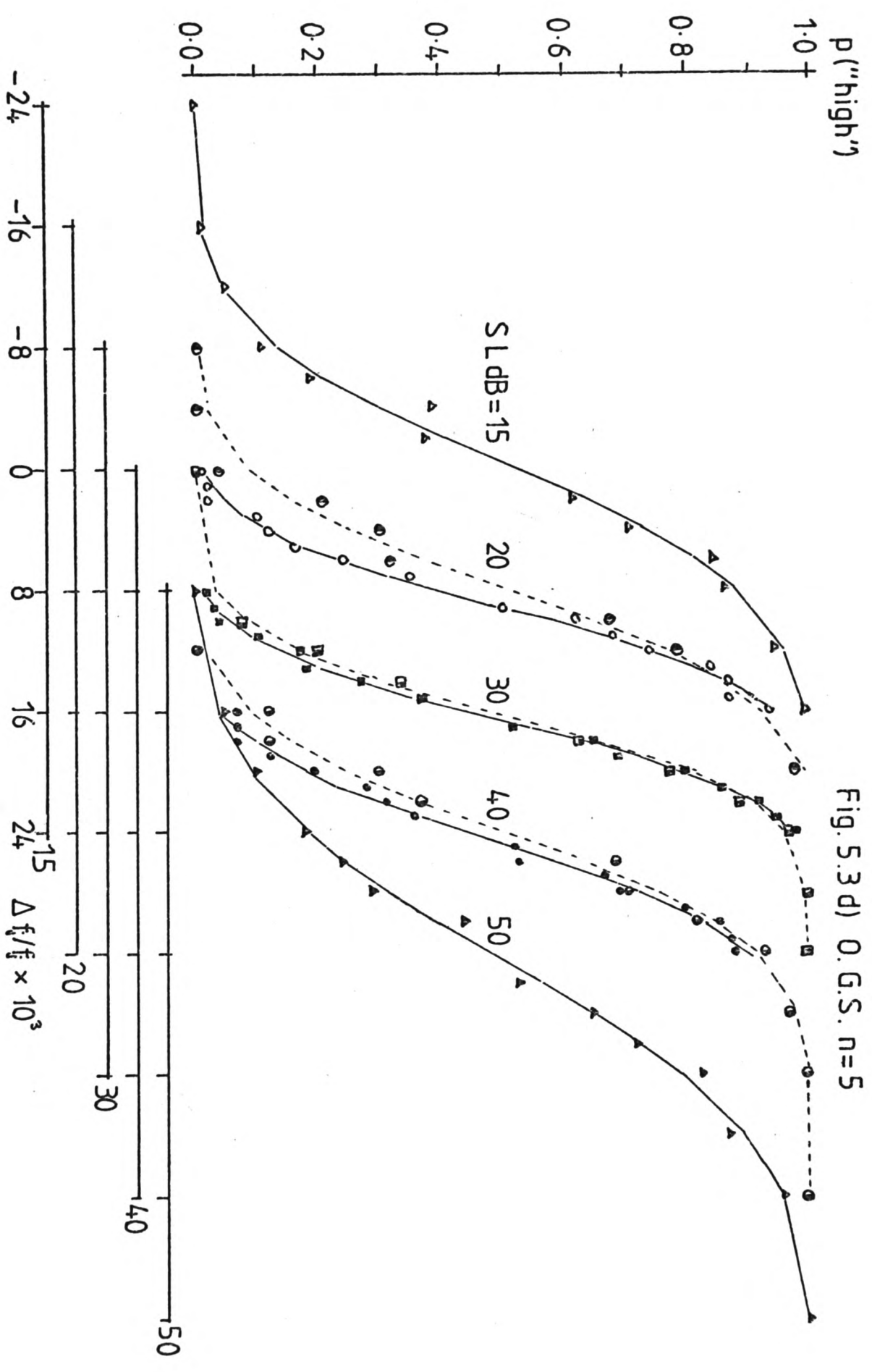
Symbol	SL dB	Trials/point.	
		A.F.	G.S.
Δ	15	100	100
\odot	20	100	100
\circ	20 (replic.)		
	$\underline{n} = 4$	100	150
	$\underline{n} = 5$	150	100
\square	30	100	100
\blacksquare	30 (replic.)		
	$\underline{n} = 4$	100	100
	$\underline{n} = 5$	100	150
\oplus	40	100	100
\bullet	40 (replic.)	150	150
\blacktriangle	50	100	100

The functions for different signal levels are displaced laterally. The solid and dashed lines are drawn between the points of the best-fitting theoretical functions from a Probit analysis. The dashed lines are the fits to the first observed functions at 20, 30, and 40 dB SL.









it is considered reasonable to conclude that the observed psychometric functions do represent cumulative normal distributions of differences.

Table 5.2 Probit analysis of observed psychometric functions from Experiment 3 - n is the harmonic number of the test signal.											
Observer	SL	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p	
A.F.	n=4	50	100	210.70	10.72	-0.0001	0.0003	7.865	10	1.0	
		40	100	317.88	17.61	+0.0001	0.0002	10.083	8	1.0	
	replic.	40	150	311.15	12.39	+0.0000	0.0001	9.822	14	1.0	
		30	100	261.75	13.79	+0.0003	0.0002	11.932	8	1.0	
	replic.	30	150	312.98	12.49	+0.0003	0.0001	18.469	14	1.0	
		20	100	113.16	6.81	+0.0012	0.0004	6.856	8	1.0	
	replic.	20	100	202.08	6.89	-0.0001	0.0002	10.256	14	1.0	
		15	100	93.54	4.55	-0.0006	0.0004	7.453	12	1.0	
	n=5	50	100	202.79	10.19	+0.0001	0.0003	11.505	12	1.0	
		40	100	297.60	16.17	+0.0003	0.0002	10.698	8	1.0	
	replic.	40	150	352.57	14.51	-0.0002	0.0001	16.760	14	1.0	
		30	100	300.01	16.53	+0.0009	0.0002	5.938	10	1.0	
	replic.	30	150	330.23	13.35	+0.0004	0.0001	14.995	8	1.0	
		20	100	237.59	12.30	-0.0001	0.0002	4.719	8	1.0	
	replic.	20	150	268.42	10.56	+0.0001	0.0002	14.857	14	1.0	
		15	100	158.12	7.69	-0.0001	0.0003	10.883	12	1.0	
G.S.	n=4	50	100	156.79	7.62	-0.0002	0.0002	10.508	12	1.0	<0.010
		40	100	172.62	17.53	+0.0010	0.0006	25.481*	6	4.247	
	replic.	40	150	219.16	17.53	+0.0003	0.0001	10.643	14	1.0	
		30	100	203.30	10.22	-0.0001	0.0003	9.549*	6	1.0	
	replic.	30	150	247.94	8.08	+0.0009	0.0001	20.830	14	1.0	
		20	100	190.15	9.45	+0.0002	0.0003	4.658*	6	1.0	
	replic.	20	150	228.73	10.65	+0.0010	0.0002	27.576	14	1.970	<0.025
		15	100	158.75	13.27	+0.0003	0.0005	17.713*	6	2.952	<0.010
	n=5	50	100	106.58	5.12	+0.0004	0.0004	4.740	12	1.0	
		40	100	167.65	8.20	+0.0000	0.0003	10.069	12	1.0	
	replic.	40	150	184.63	6.57	+0.0008	0.0002	9.631	14	1.0	
		30	100	221.46	11.29	+0.0002	0.0003	3.316	11	1.0	
	replic.	30	100	241.91	9.64	+0.0005	0.0002	3.815	14	1.0	
		20	100	170.71	8.37	-0.0002	0.0003	10.171*	6	1.0	
	replic.	20	100	228.93	9.316	+0.0011	0.0002	7.543	14	1.0	
		15	100	142.11	6.85	-0.0003	0.0003	7.637	12	1.0	

Since the two test signals x_1 and x_2 differed only by Δf , it can be assumed that the variances $\sigma^2(x'_1)$ and $\sigma^2(x'_2)$ are equal. Assuming that the covariance $r(x'_1, x'_2) \cdot \sigma(x'_1) \cdot \sigma(x'_2)$ is zero, the parameter $\sigma(f'_j)/f_j$ can be estimated as $\sigma(x'_2 - x'_1) / \sqrt{2}$, where

$\sigma(x'_2 - x'_1)$ is given by the reciprocal of the slope of the theoretical function. Figure 5.4 presents the estimates of $\sigma(f'_j)/f_j$ for $f_j = 837.6$ Hz and 1047 Hz as a function of SL for both observers; the error bars in the figure are equal to ± 1.96 standard errors of the slope of the theoretical functions, and thus represent 95% confidence intervals. For SLs of 20, 30 and 40 dB, where two functions were obtained for each observer, both estimates of $\sigma(f'_j)/f_j$ are shown in the figure.

The Estimates of $\sigma(f'_j)/f_j$.

The estimates of $\sigma(f'_j)/f_j$ at 20, 30, and 40 dB SL from the performance of observer A.F. show generally rather small changes over the two replications of the experiment with different values of Δf_j ; except for $n=4$ at 20 dB SL, where the estimate of $\sigma(f'_j)/f_j$ falls from 0.006249 to 0.003499, the estimates from the second set of data are approximately 10% smaller, and this difference is not much larger than the standard error of the estimates. The estimates from the performance of observer G.S. show a similar degree of improvement over the two replications, except at 20 dB SL, where there is a decrease in the estimate of about 15% for $n=4$ and 20% for $n=5$, and for $n=5$ at 40 dB SL, where the second estimate is about 7% larger than the first. There is no indication from the values of χ^2 that the fit of the theoretical psychometric functions is any better for the second set of data, so the differences between the two replications are attributed to the effects of practice, and it may be expected that the estimates of $\sigma(f'_j)/f_j$ at 15 and 50 dB SL, where no replication was

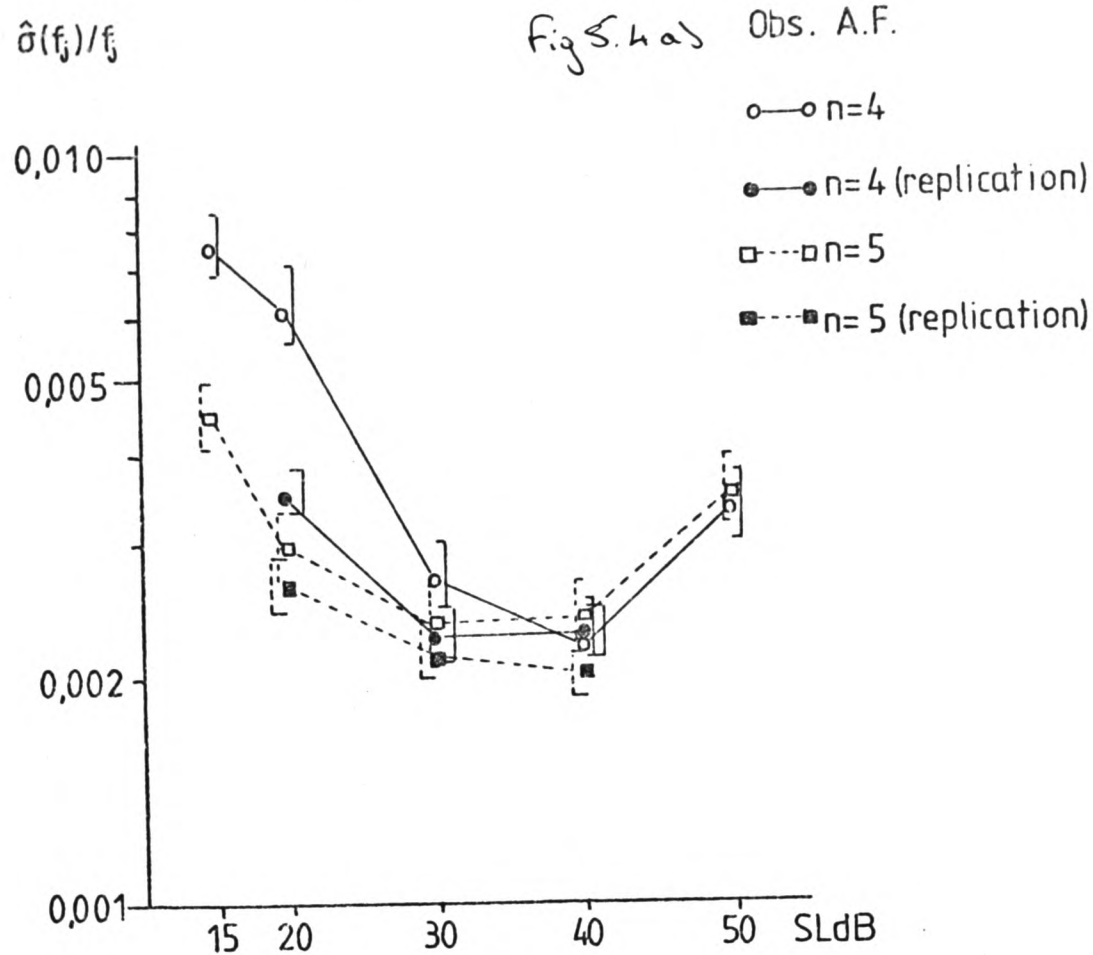


Fig. 5.4 b)

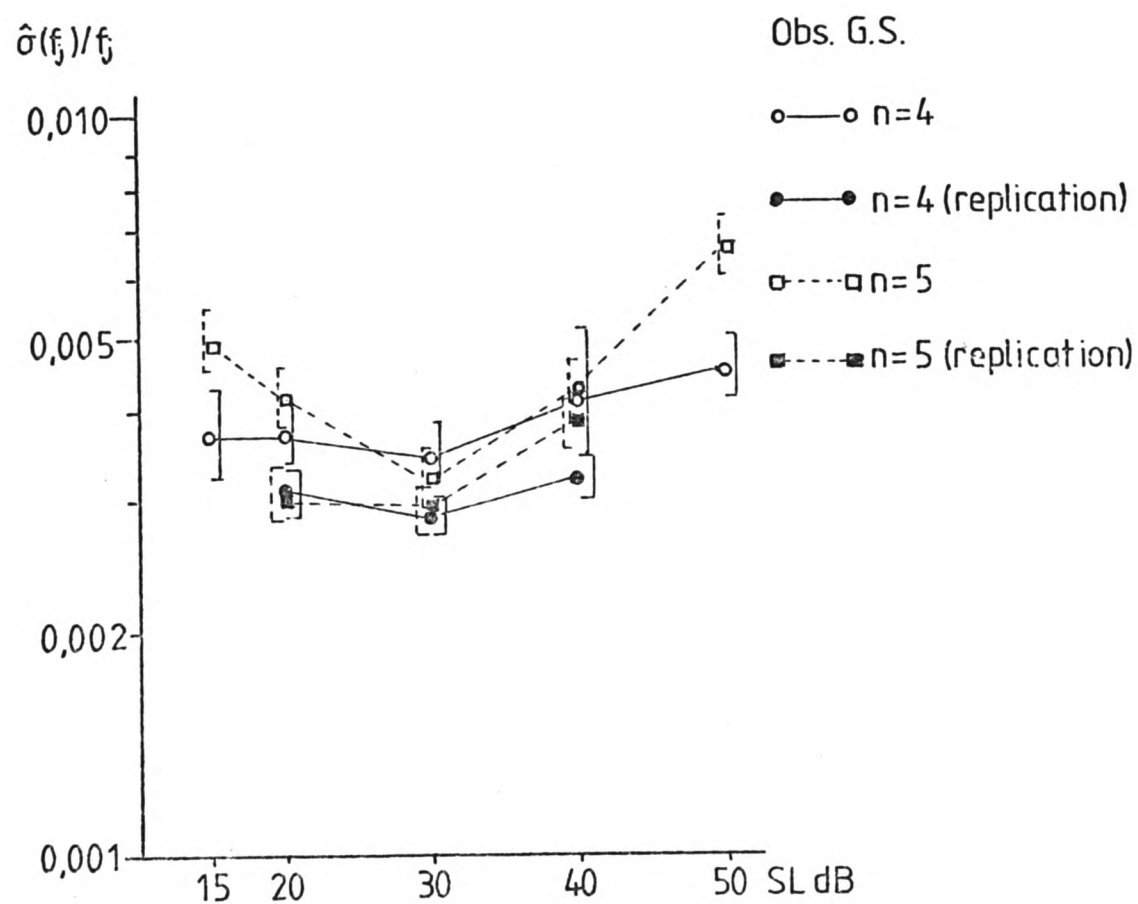


Figure 5.4. Estimates of $\hat{\sigma}(f_j)/f_j$ as a function of Sensation Level from Experiment 3. Panel a) presents the estimates from observer A.F., and panel b) the estimates from observer G.S. The error bars represent 95% confidence limits.

The legend n indicates the harmonic number of the test signal with respect to a fundamental frequency of 209.4 Hz.

performed, would have shown a similar degree of decrease after the observers had been further practised. The decrease in the estimates of $\sigma(f'_j)/f_j$ over the replications is, however, of the same order as the standard error of the estimates, and for this reason, a replication at 15 and 50 dB SL was not considered necessary.

The relationship between the presentation level of the signals and the estimated $\sigma(f'_j)/f_j$ is non-monotonic; the estimates decrease with SL up to 30 to 40 dB, and show a marked increase at 50 dB SL. Some part of this non-monotonicity is presumably due to the constant signal to noise ratio of about 40 dB in the region of the components for signal levels of 40 to 50 dB (see table 5.1), which would be expected to lead to a constant value for $\sigma(f'_j)/f_j$ at these levels. This factor is not however a sufficient explanation of the increase in the DL at 50 dB. Weir, Jesteadt, and Green's (1977) study of the DL as a function of SL for sinusoidal signals shows no such non-monotonicity; their measured DLs continued to decrease with level up to 80 dB SL. A study reported by Henning (1967b), where sinusoids were presented in a 32 dB spectrum level white noise, does however show some indication of an increase in the DL at signal levels of 60 dB above the noise. The signals used in the experiment described here were not of course single sinusoids, so it may be the case that at higher signal levels, there is some interaction between the two components or between the components and the low-pass noise.

Since masking spreads upwards in frequency (Egan and Hake,

1950; Wegel and Lane, 1924) the larger DLs at 50 db SL may simply be due to an increased masking effect from the low-pass noise. A recent study of the simultaneous masking of a tone by a noise described by Weber (1977) shows that where the noise band and 1000 Hz tone are 200 Hz or more apart, little masking is evident with noise spectral-densities up to 30 dB, but above this noise level, the masked threshold of the signal increases more rapidly than the masker level.

The values of the estimates of $\sigma(f'_j)/f_j$ from these data are somewhat larger than estimates of the DL for a single sinusoids at the same frequencies (Weir, Jesteadt, and Green, *ibid*), suggesting that the presence of either the second component or the masking noise did contribute to the estimated standard deviations. The estimates obtained here are also larger than those obtained from the same two observers for frequency discrimination of a 1200 Hz sinusoid in a silent background (see Experiment 2, Chapter 4). These estimates are also up to four times smaller than the value of 0.010 which Goldstein (1973) found appropriate for the musical interval identification data of Houtsma and Goldstein. The current estimates are also smaller than those appropriate for the estimates of $\sigma(f'_0)/f_0$ from Experiment 1 above. A possible cause of this discrepancy may be the effect of stimulus uncertainty, and this matter receives further discussion below.

If the estimates of $\sigma(f'_j)/f_j$ obtained here are appropriate as estimates of the noisiness of component frequency estimates at pitch processing, discrimination as a function of

sensation level for the residue pitch of the complex signal used here can be predicted from these estimates by the spectral-pattern theories and discrimination for the residue pitch of the signal would be expected to show a similar dependence upon the sensation level of the components as that shown by frequency discrimination for the component frequencies. If however, cues derived from the residual periodicity information carried in the auditory nerve fibres are responsible for the residue pitch of this signal, it would be expected that discrimination for the residue pitch of the signal may differ from that predicted by the spectral-pattern recognition models, and that $\sigma(f'_0)/f_0$ would be no smaller than the smaller estimate of $\sigma(f'_j)/f_j$ derived here. Experiments 5 and 6, which investigate discrimination for the residue pitch of such a two-component complex signal as a function of the presentation level of the components, are described in the following chapter, and the above predictions are examined in there.

CHAPTER 6

Abstract.

This chapter describes two experiments designed to estimate $\sigma(f'_0)/f_0$ for the residue pitch of a two-component harmonic signal as a function of the presentation level of the components of the signal. Since it has been shown above (see Chapter 2) that this pitch need not be single-valued, one of the experiments (Experiment 5) was designed so as to maximize the likelihood that the observers were indeed listening to a residue pitch in the region of the fundamental frequency of the signals. Here, the pitch comparison was made between the residue pitch of a two-component harmonic signal and the pitch of a signal composed of the 1st, 2nd, 7th, 8th, and 9th harmonics of a similar fundamental frequency close to 200 Hz. The second experiment (Experiment 6) was designed so that the stimulus uncertainty was minimal. Here, the observers were instructed to compare the residue pitches of a pair of two-component signals composed of the same two harmonics (the 4th and 5th) of slightly different fundamental frequencies close to 200 Hz.

The resulting estimates are compared with relative standard deviations predicted by the spectral-pattern recognition models

of residue pitch perception from the estimates of $\sigma(f'_j)/f_j$ obtained from Experiment 3. The results of the present experiments show the same non-monotonic dependence upon presentation level as found in Experiment 3.

The estimates of $\sigma(f'_0)/f_0$ from Experiment 6 are very close to the values predicted by the spectral-pattern models from the estimates of $\sigma(f'_j)/f_j$ from Experiment 3. The estimates of $\sigma(f'_0)/f_0$ from Experiment 5 exceed those from Experiment 6 by a factor of 1.90 for one observer, and 1.57 for the other. These factors exceed those expected from the differences in stimulus uncertainty between the two experiments, which would indicate a theoretical factor of 1.22. It is suggested that the estimates of $\sigma(f'_0)/f_0$ from Experiment 5, where the pitches compared were derived from components in different frequency regions, include some additional internal noise associated with pitch comparisons between different frequency-selective channels.

Introductory Comments.

Given estimates of $\sigma(f'_j)/f_j$ for the two components of a harmonic signal, which are available from the results of Experiment 3, it is possible to predict a value for $\sigma(f'_0)/f_0$ from the spectral-pattern recognition models described by Goldstein (1973) and Wightman (1973b). As was shown in Chapter 2, if $\sigma(f'_j)/f_j$ represents the relative frequency standard deviation of the estimated spectrum in the region of f_j , then for the spectral-pattern recognition models, $\sigma(f'_0)/f_0$, the relative standard deviation of the mode of the residue pitch pdf at the estimated fundamental frequency f'_0 is given by

$$1 / [\sigma(f'_0)/f_0]^2 = \sum_{j=1}^2 \{ 1 / [\sigma(f'_j)/f_j]^2 \},$$

so that $\sigma(f'_0)/f_0$ is expected to be smaller than the smaller value of $\sigma(f'_j)/f_j$. The periodicity theory leads to the prediction that $\sigma(f'_0)/f_0$ should be greater than the smaller value of $\sigma(f'_j)/f_j$ for well-resolved components.

Experiment 1, which was described in Chapter 4, investigated pitch discrimination between the residue pitches of two two-component harmonic signals as a function of the presentation level of the components of the signals. For several reasons, this experiment did not provide results comparable with the results of Experiment 3. A serious difficulty with the data of Experiment 1 was the tendency for the observers to base their responses upon differences between

the frequencies of the components of the two signals presented on each trial. Since the component frequency differences were uncorrelated with the fundamental frequency differences between the two signals, this tendency is likely to have led to an overestimation of the DL. A further difficulty arises from the different presentation conditions employed in the two experiments; although Experiment 3 was originally intended to mimic as closely as possible the conditions of Experiment 1, it was found necessary to introduce a low-pass filtered masking noise to eliminate what were assumed to be the effects of a cubic-difference tone generated by an auditory non-linearity. Since Experiment 1 did not employ this masking noise, the results of that experiment may be further contaminated by the effects of the same non-linearity. Furthermore, this masking noise may have contributed to the estimates of $\sigma(f'_j)/f_j$, and the estimates of $\sigma(f'_0)/f_0$ from Experiment 1 do not include such a contribution. Two experiments (Experiment 5 and 6) intended to estimate $\sigma(f'_0)/f_0$ are described here. Since the task differs from that used in Experiment 3, it is impossible to replicate simultaneously all of the presentation conditions of that experiment, and the experiments described here represent two attempts to mimic as closely as possible some of the conditions which are considered to be of importance.

Comments on Method for Experiment 5.

Experiment 5 was designed to maximize the likelihood that the observers were listening to a residue pitch in the region of the fundamental frequency of the signal. This was achieved with

the aid of a reference signal, which was always the first of the two signals presented on each trial. The reference signal always had the same fundamental frequency, and was constructed so as to have a clear and unambiguous pitch. The second signal presented on each trial, which here is called the comparison signal, contained two harmonics. The experiment was intended to estimate the standard deviation of the internal representation of the residue pitch of this latter signal. Since the harmonic structures and presentation levels of these two signals differed, it is not reasonable to assume that their respective internal representations would have equal variances. For this reason, the decision model described in Chapter 3 requires that the variance of the internal representation of the residue pitch of the second signal must be found by subtraction of the variance of the internal representation of the reference signal from the variance of the distribution of differences. It is therefore necessary that the variance of the internal representation of the reference signal be independently known. This variance can be estimated from a 2AFC pitch discrimination task where two signals of the form of this reference signal, which differ only in their fundamental frequency, are presented in the two temporal intervals. Experiment 4 was carried out for this purpose, and is described below.

In order to avoid the problems caused by the tendency observed in Experiment 1 for the observers to judge the relative pitches of two complex signals by reference to differences in their respective component frequencies, the reference signal was constructed so as to contain components both higher and lower in

frequency than those of the comparison signal. The reference signal employed here contained a component at its fundamental frequency of 209.4 Hz, and components at the 2nd, 7th, 8th and 9th harmonics of this fundamental frequency. Because this signal included the fundamental frequency and several harmonics, its pitch was judged to be clear and unambiguous by the two observers. The harmonics included were selected so that they would have different harmonic numbers from either the two components of the comparison signals or any cubic-difference-tone which might be generated from those components by an auditory non-linearity. The reference signal was always presented in the first interval of each trial, and also served to inform the observer of the pitch region to which he should attend when judging the pitch of the comparison signal. The fundamental frequency difference between the two signals was always small with respect to the differences between the means of the modes of the theoretical pdf of the residue pitch of the comparison signal, so that this stimulus configuration allowed the estimated standard deviation of the internal representation of the pitch of the comparison signal to be interpreted with some confidence as the standard deviation of the mode of the residue pitch pdf having its mean at approximately the fundamental frequency of the comparison signal.

One problem which might arise with this experimental method is as follows. Since the reference signal is of constant fundamental frequency, the comparison signals have fundamental frequencies Δf above or below that of the reference signal, and

$|\Delta f|$ is constant during a block of trials, the fundamental frequency differences between the comparison signals presented on successive trials will either be equal or will differ by $2\Delta f$. If the harmonic numbers of the comparison signals are always the same, observers might be able to utilize differences between either the residue pitches or the component frequencies of the comparison signals presented on successive trials as a cue. Such a strategy would lead to responses which would be indistinguishable from responses based upon a comparison of the pitch of the reference signal with the residue pitch of the comparison signal given on the same trial. In order to discourage the use of such a strategy, each block of trials included two comparison signals having different pairs of harmonic numbers. To facilitate a direct comparison with the signal employed in Experiment 3, one of these comparison signals was composed of the 4th and 5th harmonics of $209.4 \pm \Delta f$ Hz, while the second signal was composed of the 5th and 6th harmonics of the same fundamental frequency. The inclusion of these two comparison signals does not prevent observers from making such between-trial comparisons, but it is possible to subject such a hypothesis to an experimental test. Experiment 11, which is described in Chapter 7, was intended for this purpose. In that experiment, different values of the fundamental frequency difference $|\Delta f|$ were presented within the same block of trials. Since the results of that experiment are similar with those for a similar experiment (Experiment 8) where different values of $|\Delta f|$ were presented in different blocks, this possibility can be rejected.

The Effect of Frequency Uncertainty.

While the experimental method described immediately above has a number of advantages, as a method for estimating $\sigma(f'_0)/f_0$, it has one major disadvantage. The differences between the harmonic numbers of the components of the comparison signals presented on different trials require the observer to attend to three different frequency regions, while on any given trial, a component will only occur in two of these regions, so that there is a degree of stimulus uncertainty. According to the multiple-band model of frequency discrimination described in Chapter 3, if an observer were to base an estimate of the fundamental frequency of the complex signal on the information derived from \underline{m} independent channels, each being subject to a noise process, and \underline{n} of these channels received a signal, the signal information would be summed over \underline{n} channels, and the noise over all \underline{m} channels. If $\underline{n} < \underline{m}$, the noise processes from the $\underline{m} - \underline{n}$ channels receiving no signal will be included in the sum of the information upon which the decision is based, leading to a more variable estimate of fundamental frequency than that derived from a process which monitored only \underline{n} channels.

Comments on Method for Experiment 6.

This experiment was designed to eliminate frequency uncertainty; here the observer had only to monitor two frequency-selective channels in order to estimate the frequencies of the two components of the signal. All of the signals presented throughout the experiment contained the same

pair of harmonics, and throughout a block of trials, the signals presented differed only by a constant fundamental frequency difference Δf . Because the two signals presented on a given trial were so similar, the design of this second experiment does not ensure that the observer is comparing the residue pitches of the two signals, and it is possible to identify two other possible strategies based upon the use of cues which are completely correlated with the difference between the fundamental frequencies of the two signals.

The observer could simply compare the frequencies of one of the harmonics common to the two signals. If an observer adopted this strategy, and if the estimates of the relative frequency difference limen for one component of the equivalent two-component complex derived from Experiment 3 are accurate, then the observer's relative DL in this experiment should be no smaller than the smaller of his relative DLs for the two components as measured at the same signal levels in Experiment 3. The possibility that this strategy is used can be rejected if this relation does not hold in the observed data. It should be noted that this relation is also predicted by the periodicity theory (see Chapter 2).

Alternatively, an observer may compare simultaneously the frequency differences between both of the components of the two signals, and take no account of the residue pitch of the signals. Since these frequency differences are correlated, it would be expected that his relative DL for the two-component complex would be smaller than the smaller of the two relative

DLs for those components. Regardless of whether frequency is coded temporally or by the amplitude or energy passed by a band-pass filter, the multiple-band model described in Chapter 3 allows the prediction of the size of this difference. The response in a frequency discrimination task for a harmonic complex signal, will be based a likelihood ratio whose denominator is the summed means of the distributions of evidence for some frequency difference, and whose divisor is the ^{square root of the} summed variance of those distributions of evidence. For a signal of two components whose frequencies are approximately known by the observer, and whose internal representations have equal variance on a scale of relative frequency, this likelihood ratio, and hence the relative DL, will thus be smaller than that derived from a single component by a factor of $\sqrt{2}$.

If then, an observer adopts this latter strategy of simultaneously comparing the frequencies of both of the components, the expected relative DL is the same as that predicted for the relative DL of the residue pitch of the signals by the spectral-pattern recognition models. If the relative DLs derived from this experiment are smaller than the smaller of the two DLs for the components as estimated from Experiment 3, it will thus be impossible to reject either of these two explanations of that result. This convergence of predictions arises simply because the initial stages of the spectral-pattern recognition models are identical to the initial stages of a multiple-band frequency coding system, and the observer's judgements will reflect either the standard deviation of his internal representation of the residue pitch of the

signal, or the equivalent standard deviation of the ensemble of cues from which that pitch may be assumed to be derived. The method adopted for Experiment 5 avoids this simple convergence because the components of the two signals presented on each trial differ greatly in frequency, and a joint interpretation of the results of these two tasks should provide a test of the possibility that the task of Experiment 5 is performed without reference to the residue pitch of the signal.

Observers.

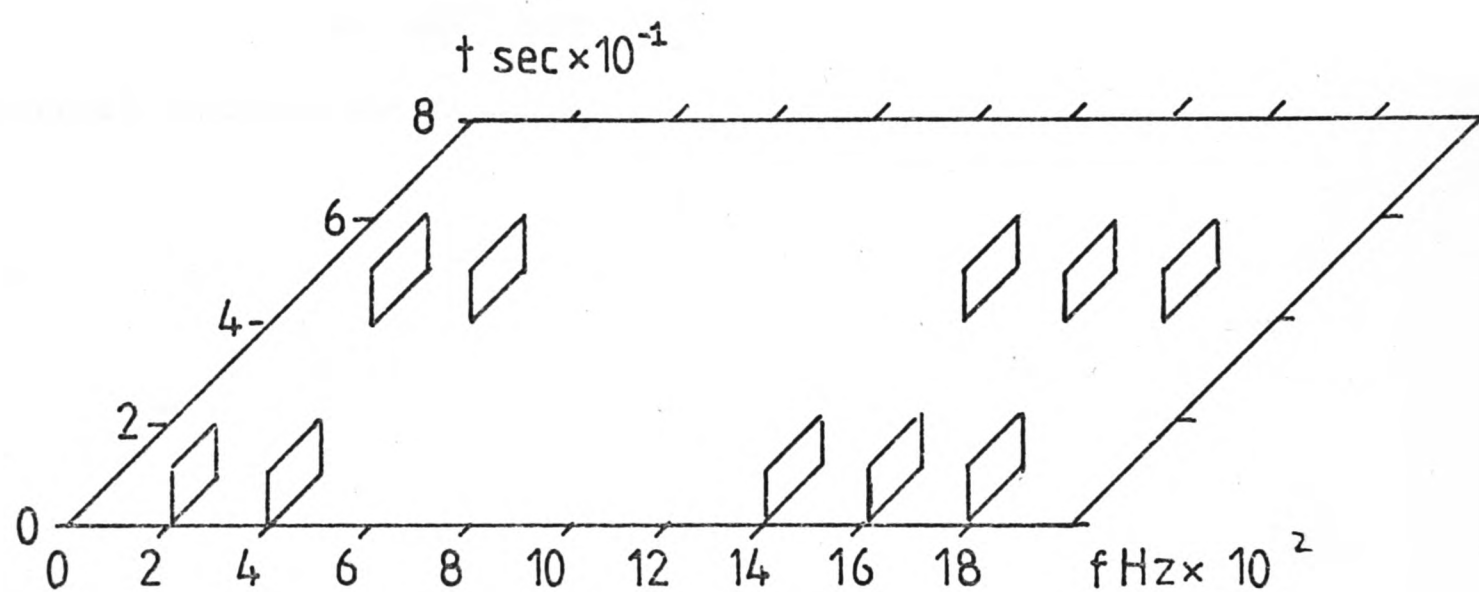
The two observers, A.F. and G.S., who participated in these experiments had each also participated in the experiments described in Chapters 4 and 5.

Experiment 4.

Stimuli.

Both of the signals presented on a trial were similar complex harmonic signals containing the 1st, 2nd, 7th, 8th and 9th harmonics of $209.4 + \Delta f/2$ or $209.4 - \Delta f/2$ Hz; the results of a spectral analysis of the presented signal appear in table 4.2. This signal is that employed as a reference signal in Experiment 5. All components were present at the same sound pressure level, this being 40 dB above the observer's threshold at 1047 Hz with a silent background. The experiment was run without background noise. The time course of a trial is illustrated in figure 6.1.

Figure 6.1. The time-course of a single trial for Experiment 4,
Both signals have the form of the reference signal for Experiment 5.



Procedure.

The absolute value of Δf was constant throughout a block of 50 trials, in each of which, the probability that the second signal had the higher fundamental frequency was 0.50. The absolute value of Δf ranged from 209.4×0.01218 Hz to 209.4×0.00101 Hz. Two series of eight blocks were presented in an experimental session of approximately one hour in length, one series being given with the absolute value of Δf increasing over blocks, and one with that value decreasing. Observer A.F. received 200 test trials at each absolute value of Δf , while observer G.S. received 400 test trials at each value. No trial-by-trial feedback was provided.

Experiment 5.

Stimuli.

The first signal presented on each trial was the reference signal presented in Experiment 4, except that the fundamental frequency of this signal was always equal to 209.4 Hz. Each component of this signal was presented at a level 40 dB above the observer's threshold at 1047 Hz in a silent background. During the presentation of the reference signal, no background noise was present. The second signal, the comparison signal, contained either the 4th and the 5th ($n=4$) or the 5th and the 6th ($n=5$) harmonics of $209.4 \pm \Delta f$ Hz; each of the two components was presented at the same sound pressure level, this being, in different series of blocks, 15, 20, 30, 40 or 50 dB

above the observer's threshold at 1047 Hz with a silent background. 100 msec after the offset of the reference signal, a low-pass filtered noise signal was gated on with a rise-time of 25 msec, and this noise remained on for 100 msec after the offset of the comparison signal. The noise was low-pass filtered at 613 Hz with a filter slope of approximately 90 dB/octave. The noise spectral-density at 613 Hz was approximately 15 dB below the level of the two components of the comparison signal as measured with a wave analyser (HP 3580A) set for a 1 Hz measuring bandwidth. The measured levels of the noise and comparison signal components at the various signal levels are given in table 6.1. The low-pass noise was introduced in order to mask any cubic-distortion products generated by the comparison signal having $n=4$.

The sequence of stimuli within a trial is illustrated in figure 6.2.

Table 6.1. Signal level and noise spectral-density (N_o) in dB relative to a 0 dB S.L. sinusoid at 1047 Hz.		
Signal level	N_o between 837 and 1047 Hz	N_o at 618 Hz
+50	+3	+35
+40	-1	+26
+30	-8	+12
+20	-10	+3
+15	-9	-5

Signal and noise levels were measured electrically at the headphone input with a Hewlett-Packard 3580A wave analyser and a bandwidth of 1 Hz. The equipment used for presenting the signals had a noise floor at about -9 dB relative to the reference level of 0 dB S.L. at 1047 Hz.

Figure 6.2. The time-course of a single trial for Experiment 5. The first signal, the reference, is presented in a silent background, and a low-pass filtered noise is present together with the second signal, the comparison.

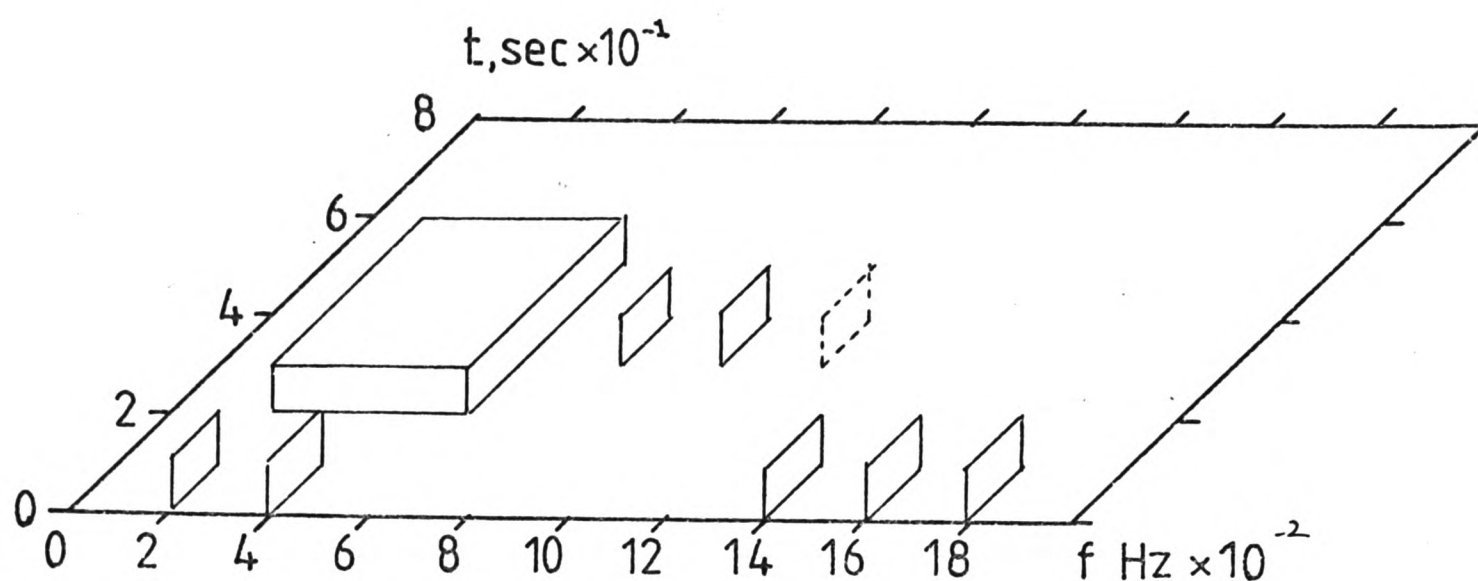
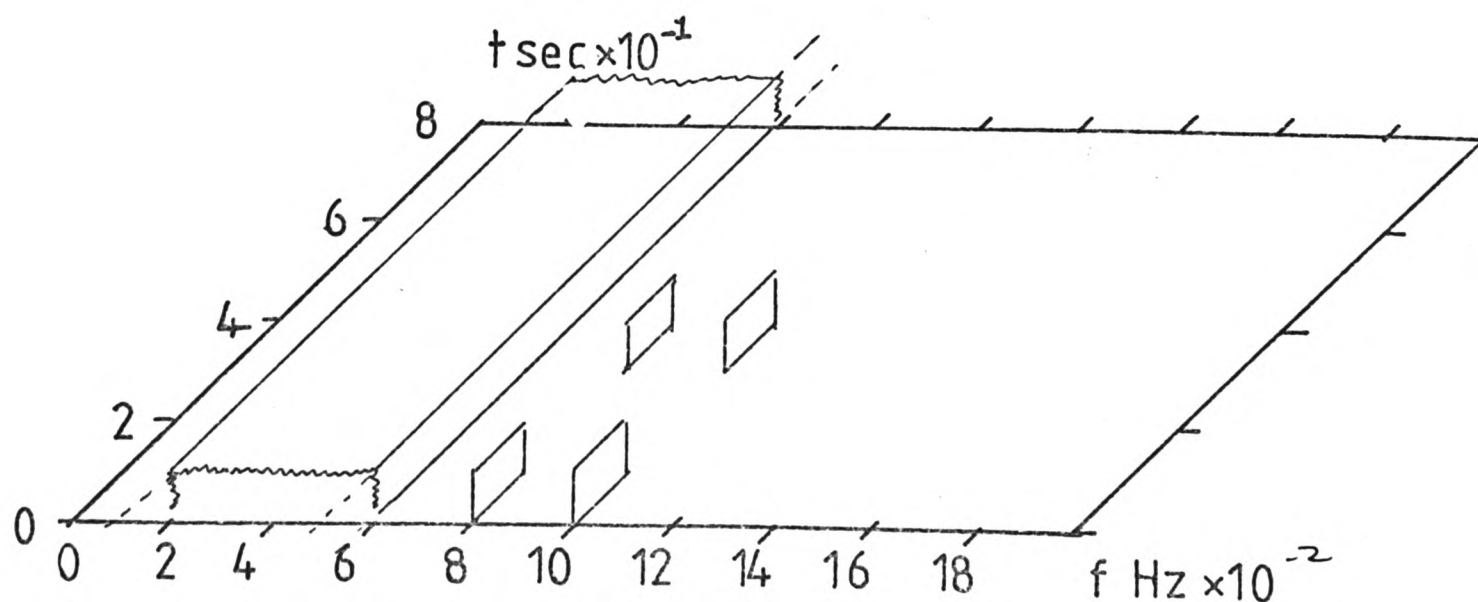


Figure 6.3. The time-course of a single trial for Experiment 6. The low-pass noise is present continuously.



Procedure.

Trials were presented in blocks of 100 with $|\Delta f|$ constant. The values of $|\Delta f|$ ranged from 209.4×0.03248 Hz to 209.4×0.00101 Hz. Within a block, the probability that the comparison signal had a higher fundamental frequency than the reference signal was 0.50, the probability that the comparison signal contained the 4th and 5th harmonics was 0.50, and these two events were independent. A series of eight blocks with a constant presentation level for the comparison signal was presented in a test session of about one hour in length. Two stimulus tapes were prepared with $|\Delta f|$ either increasing or decreasing over blocks. The order of test sessions at the different comparison signal levels and ordering of $|\Delta f|$ was approximately counterbalanced over the days of testing. Each observer received at least four series of blocks at each presentation level, and thus received at least 200 trials for each of the two comparison signals at each absolute value of Δf and each presentation level.

Experiment 6.

Stimuli.

Both of the stimuli presented on each trial contained the 4th and 5th harmonics of $209.4 + \Delta f/2$ or $209.4 - \Delta f/2$ Hz. The two components of each signal were presented at the same sound pressure level, which, in different series of blocks was either 15, 20, 30, 40 or 50 dB above the observers threshold in quiet

at 1047 Hz. These signals were presented against the same low-pass filtered background noise as that presented in Experiment 5, except that this noise was on continuously. The events in a trial are illustrated in figure 6.3.

Procedure.

Each block contained 50 trials, and the probability that the second signal had the higher fundamental frequency was 0.50. The values of $|\Delta f|$ ranged from 209.4×0.01218 Hz to 209.4×0.00101 Hz. Each test session lasted about an hour, and comprised two series of eight blocks, the presentation level being constant within each series. $|\Delta f|$ either increased or decreased over the blocks of a series, and the ordering over test sessions of blocks at different signal levels and with $|\Delta f|$ increasing or decreasing was approximately counterbalanced. Each observer received at least 200 trials at each absolute value of Δf and at each presentation level. Two stimulus tapes were prepared, each being presented equally often, and played both forwards and backwards.

Results.

Psychometric functions.

Figures 6.4, 6.5, and 6.6 present the observed psychometric functions for Experiments 4, 5 and 6 respectively, together with the best-fitting theoretical functions derived from a Probit analysis of the data. The parameters of the theoretical

Figure 6.4. The psychometric functions from Experiment 4. Pitch discrimination for the reference signal. The functions for the two observers are laterally displaced. The solid line is drawn between the points of the best-fitting theoretical function from a Probit analysis. The number of trials per point is 100 for observer A.F., and 200 for observer G.S.

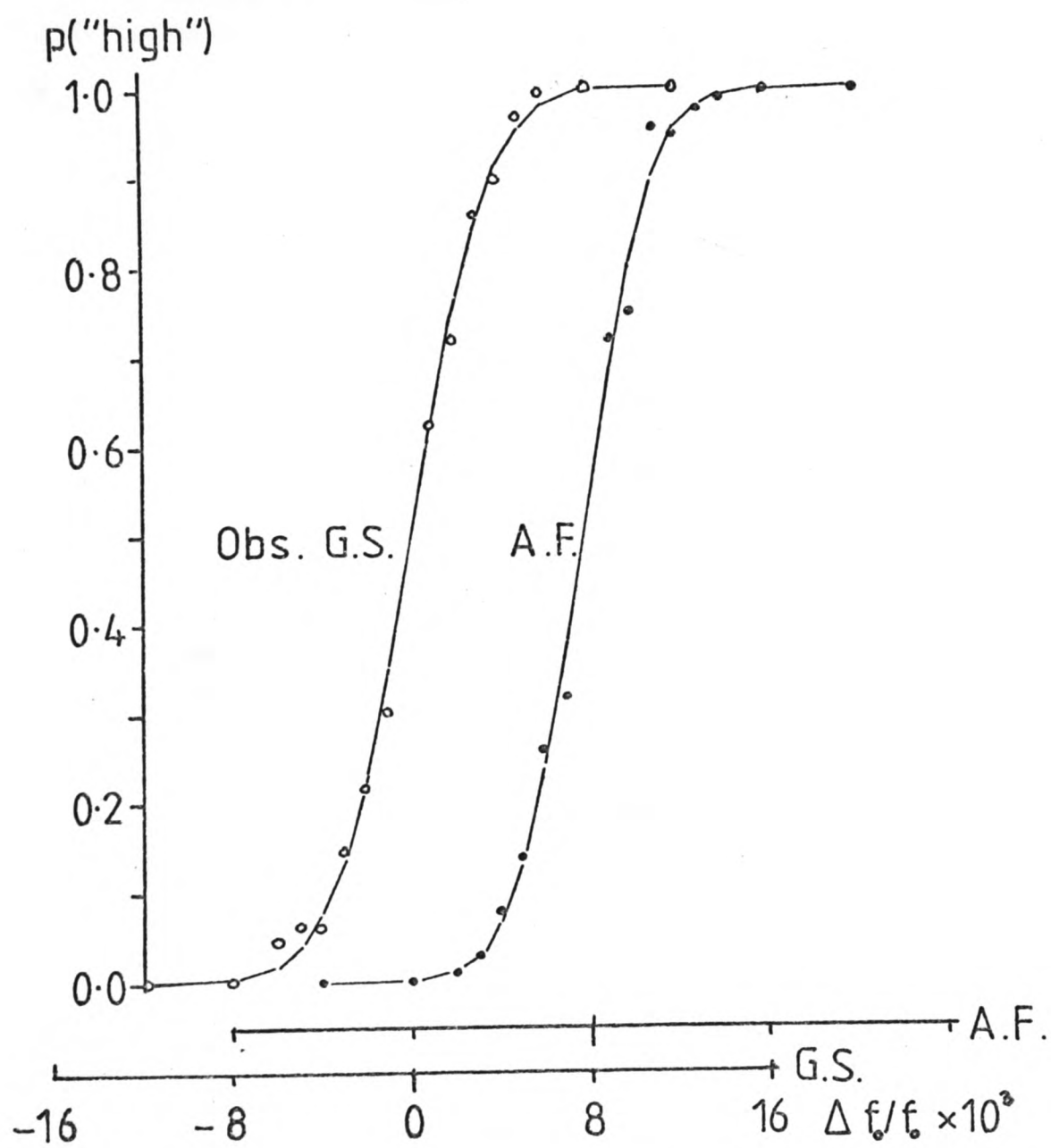


Figure 6.5. Psychometric functions from Experiment 5; pitch discrimination between the reference signal and a two-component comparison signal. The functions for different presentation levels of the comparison signal are laterally displaced. The lines are drawn through the best-fitting theoretical function.

Panel a) observer A.F., $\underline{n} = 4$.

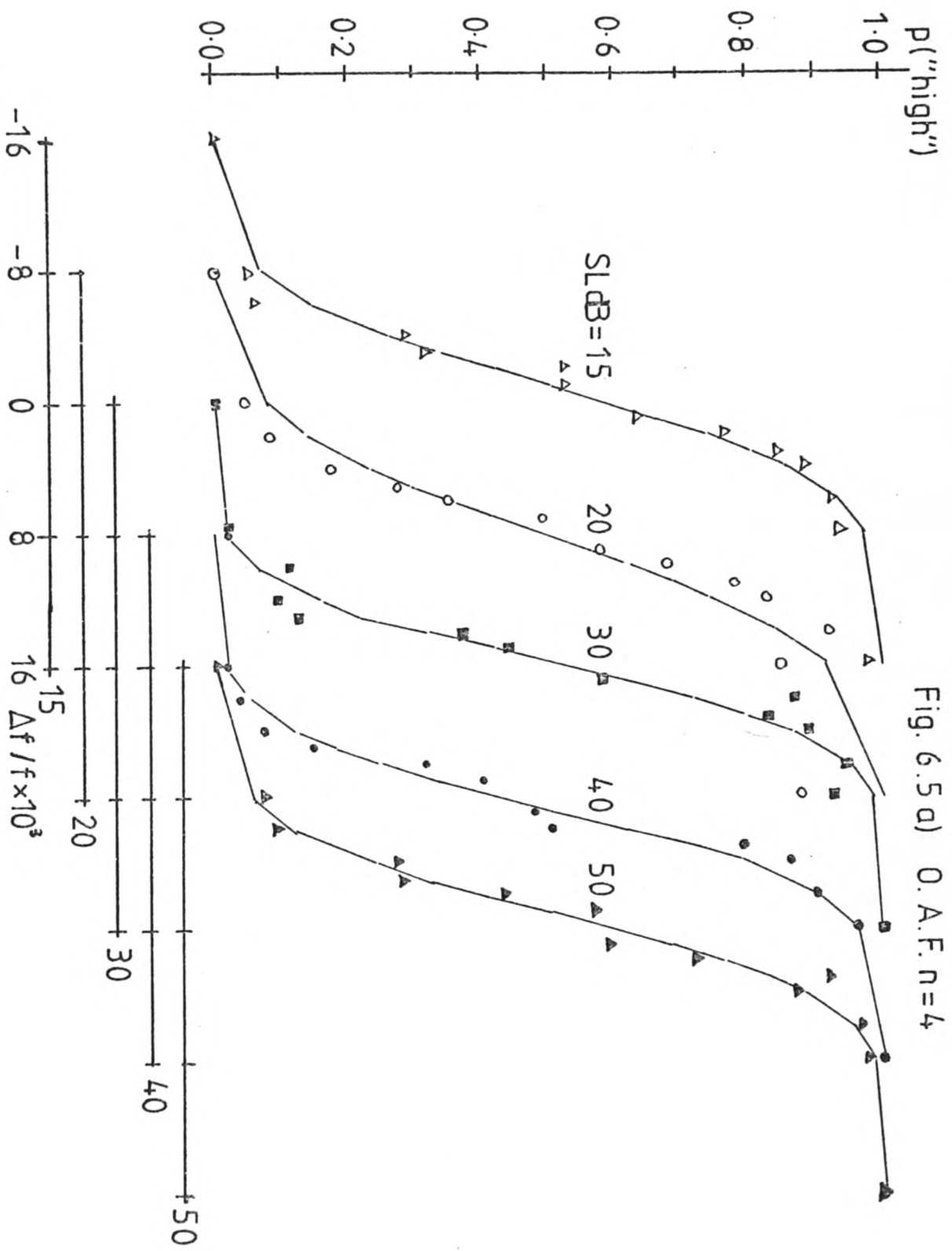
Panel b) observer A.F., $\underline{n} = 5$.

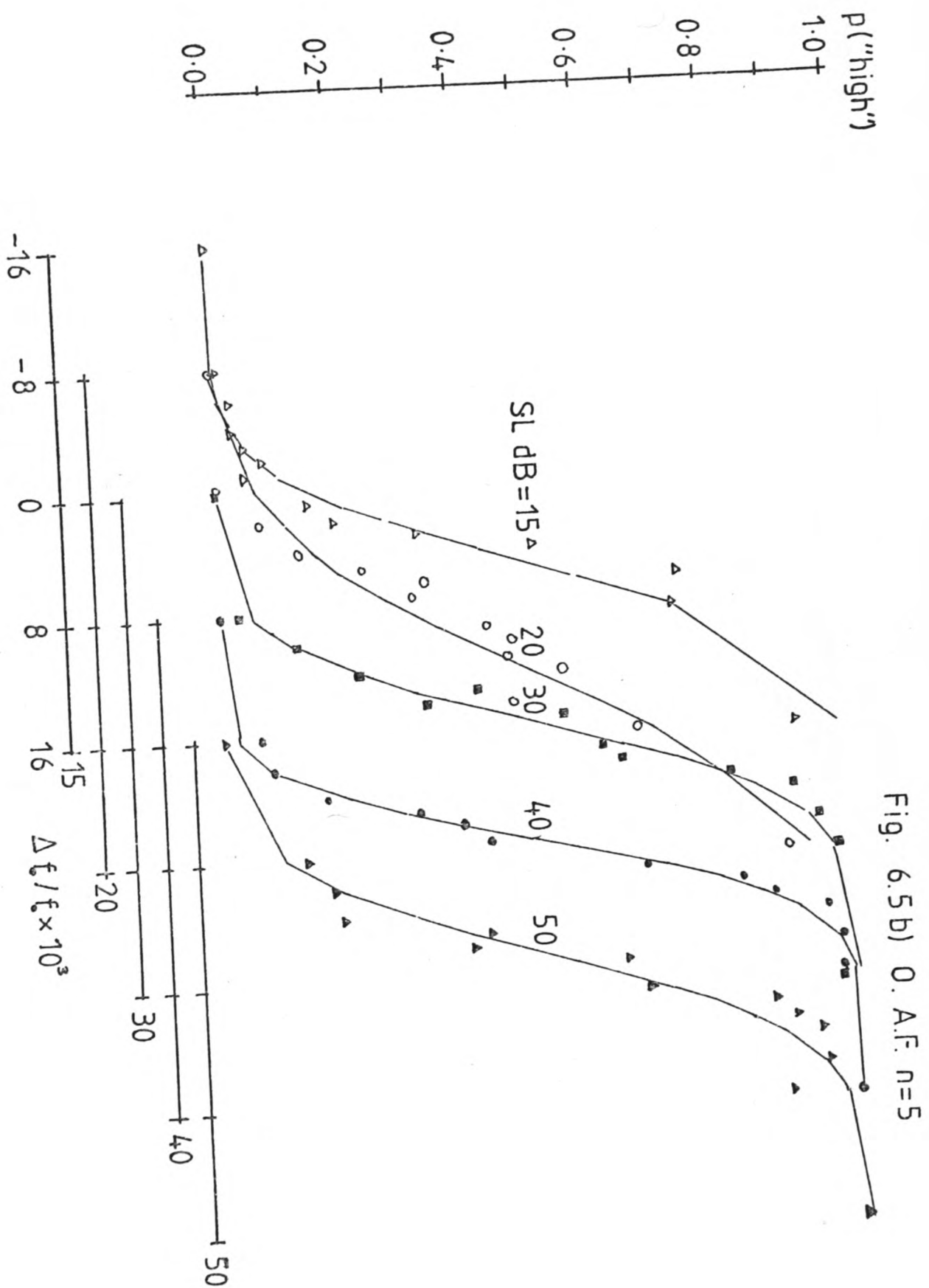
Panel c) observer G.S., $\underline{n} = 4$.

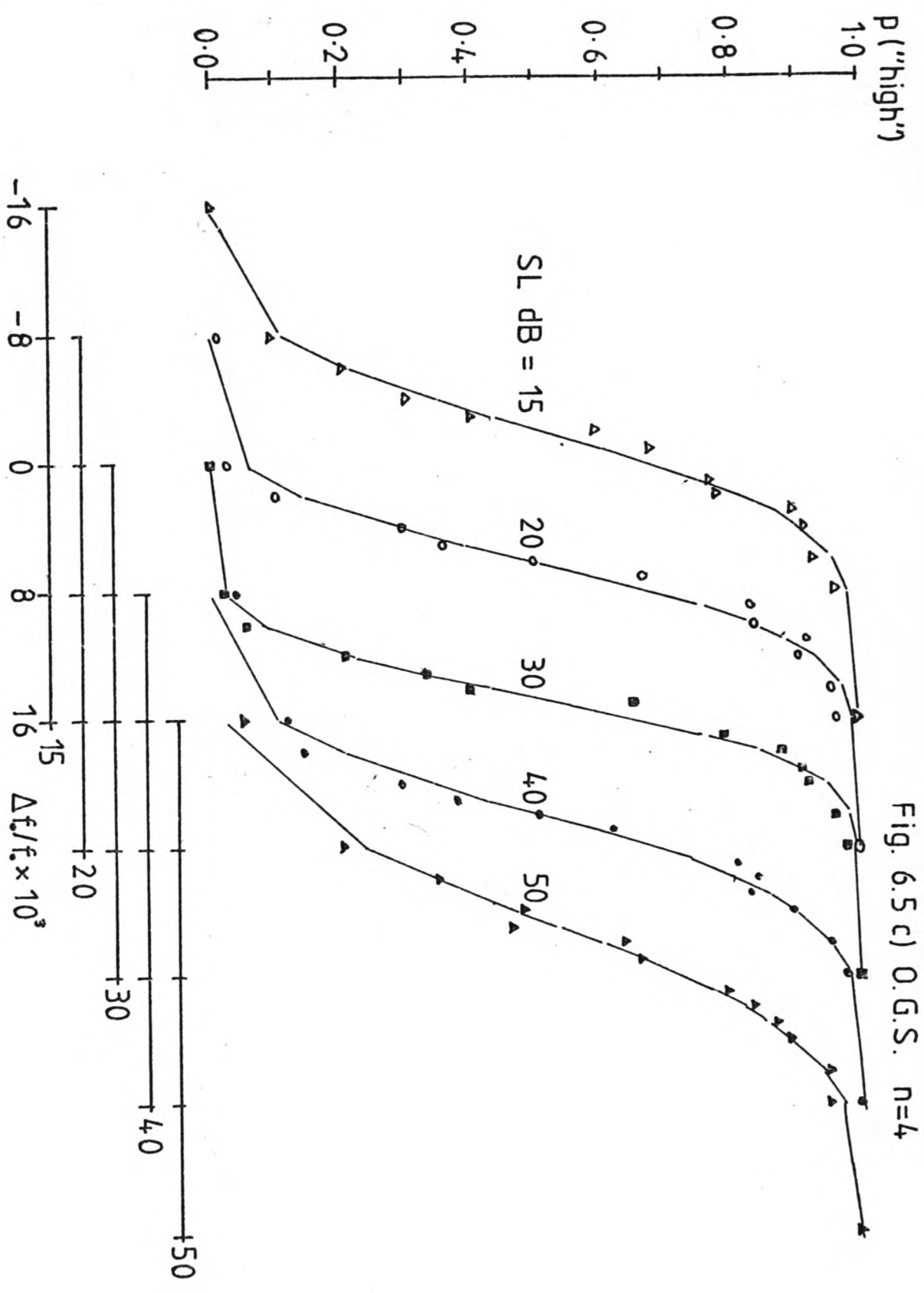
Panel d) observer G.S., $\underline{n} = 5$.

The legend \underline{n} indicates the lower harmonic number of the comparison signal.

The number of trials per point is 100 for observer A.F. except at 40 dB SL, where it is 150. All points for observer G.S. are based on 150 trials.







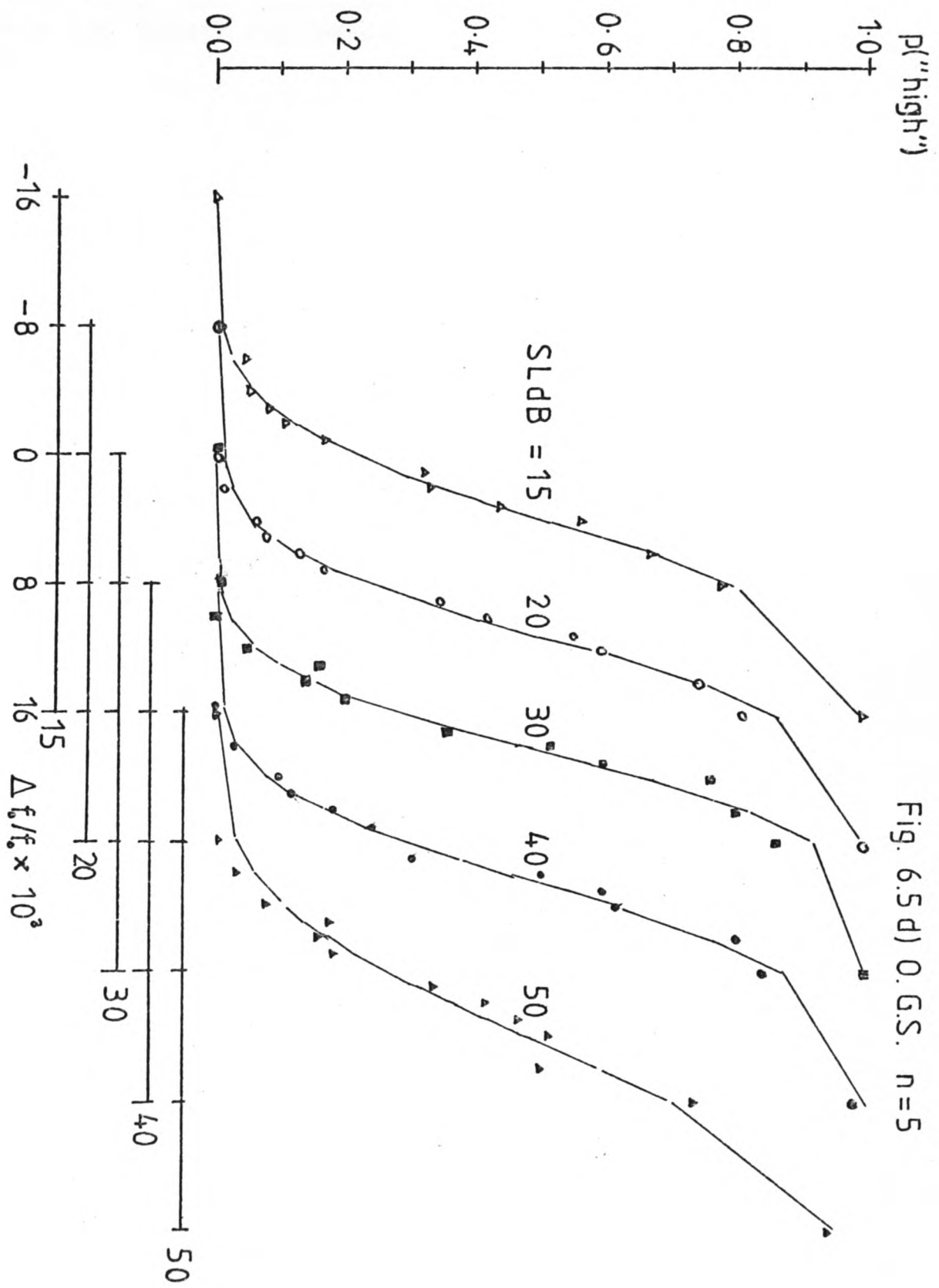
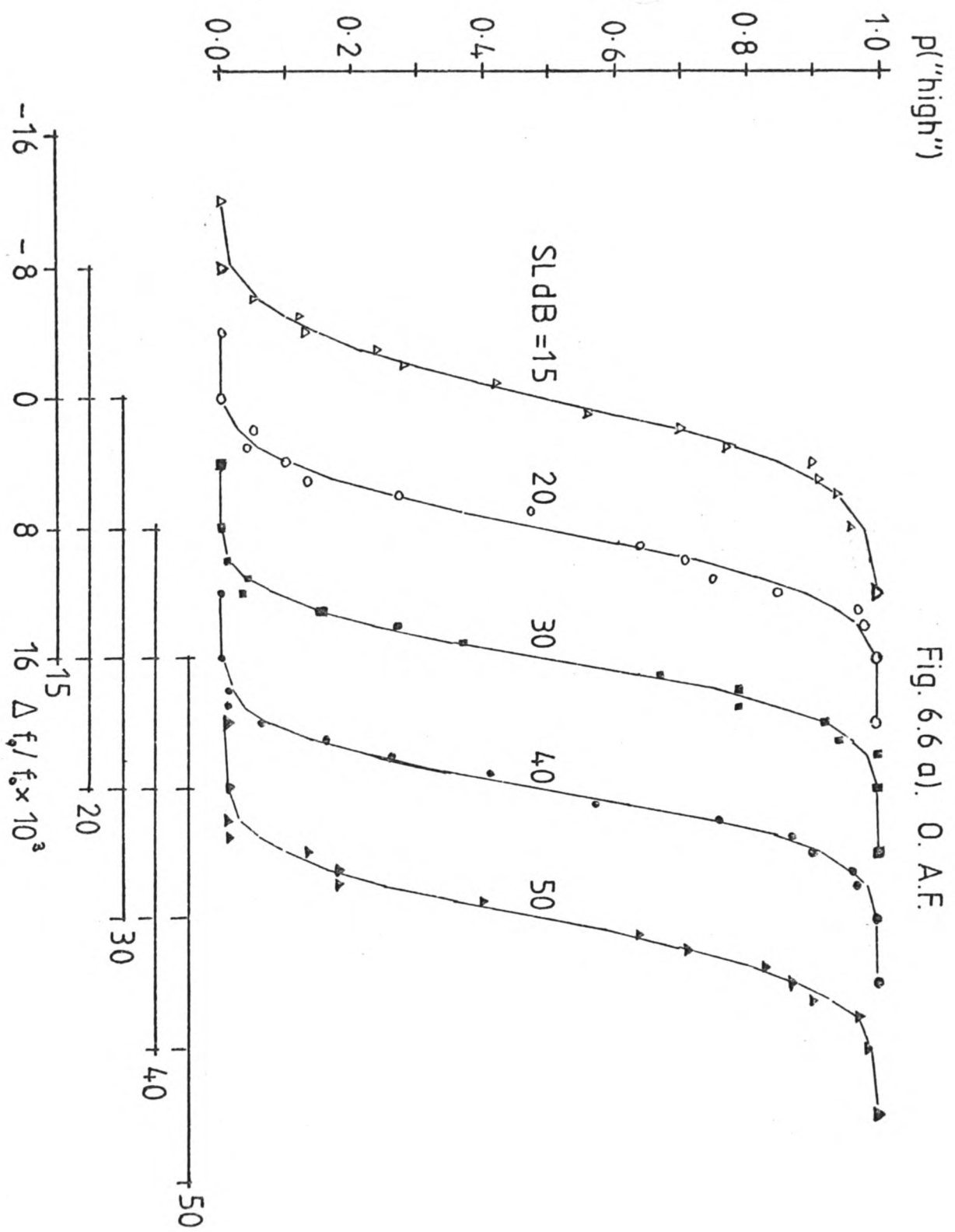
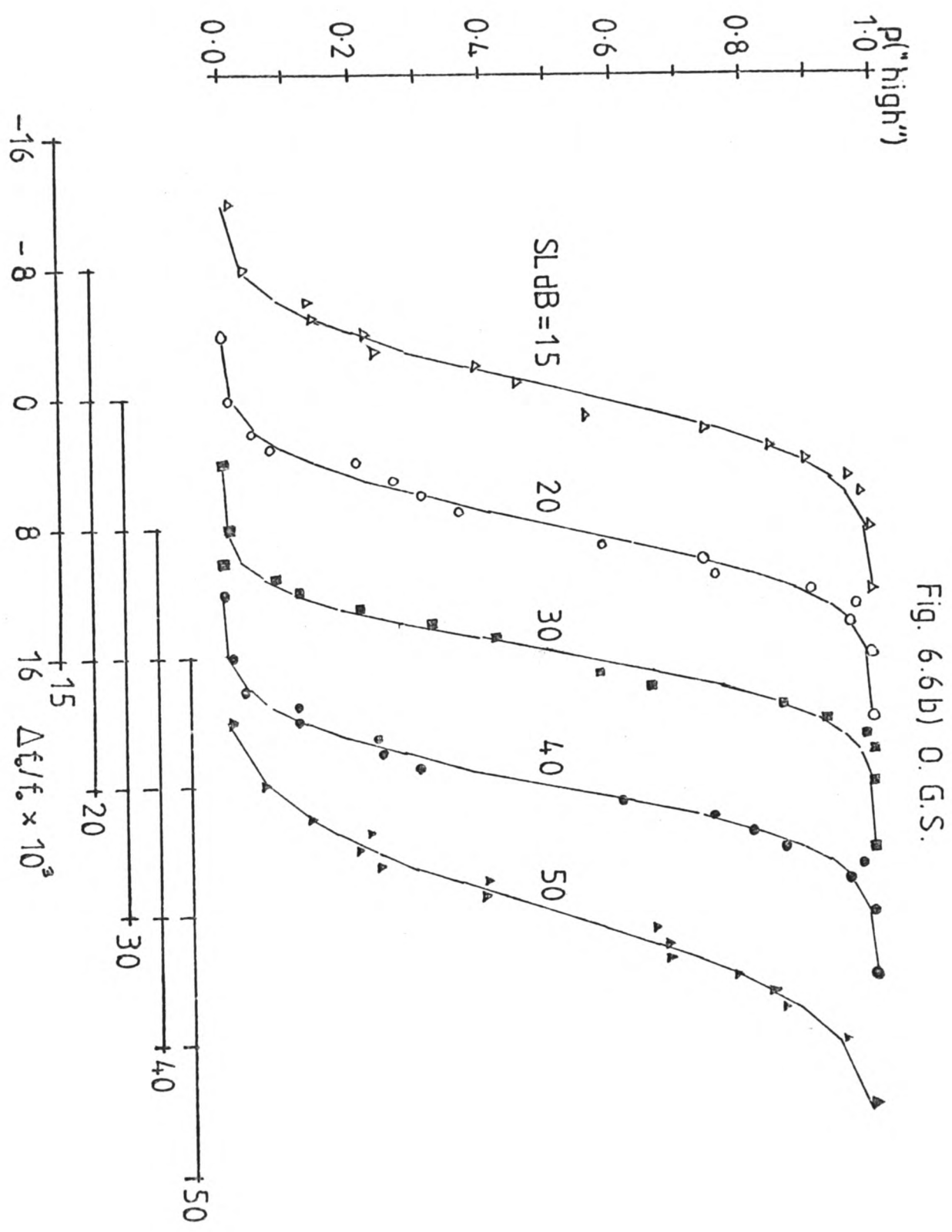


Figure 6.6. Psychometric functions from Experiment 6; pitch discrimination for a two-component harmonic complex signal containing the 4th and 5th harmonics of a frequency around 200 Hz. The solid lines are drawn between the best-fitting theoretical functions.

Panel a) observer A.F., 100 trials per point.

Panel b) observer G.S., 150 trials per point except at 30 dB SL, where there were 150 trials per point.





functions and the statistic are presented in tables 6.2, 6.3, and 6.4. In each case, the function analysed was that for the probability of the response that the second signal was higher in pitch as a function of the signed value of $\Delta f / f_0$.

Table 6.2. Probit analysis of observed psychometric functions from Experiment 4.										
Observer	SL	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
A.F.	40	100	390.04	17.76	-0.0002	0.0001	9.288	14	1.0	
G.S.	40	200	341.26	10.64	+0.0001	0.0001	20.47	14	1.0	

Table 6.3. Probit analysis of observed psychometric functions from Experiment 5 - n is the lower harmonic number of the comparison signal. Observer A.F.										
	SL	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	50	100	234.54	11.80	-0.0011	0.0002	16.923	13	1.0	
	40	150	243.12	18.65	+0.0008	0.0003	32.084*	9	3.565	<0.01
	30	100	258.67	24.50	-0.0002	0.0003	48.392	13	3.722	<0.01
	20	100	171.97	19.27	+0.0002	0.0005	41.752*	10	4.175	<0.01
	15	100	210.59	15.08	+0.0011	0.0003	17.318*	9	1.924	<0.05
n=5	50	100	203.84	21.69	-0.0017	0.0004	36.842*	9	4.094	<0.01
	40	150	273.66	11.00	-0.0010	0.0001	11.310*	7	1.0	
	30	100	219.56	11.22	-0.0003	0.0002	15.117*	9	1.0	
	20	100	121.61	11.46	+0.0004	0.0005	30.154	14	2.154	<0.01
	15	100	158.12	7.69	-0.0001	0.0003	10.883	12	1.0	

Individual Sessions at 20 dB

n=4

Session

	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
1	25	140.98	15.61	+0.0002	0.0005	6.946	14	1.0	
2	25	233.40	23.64	+0.0012	0.0004	14.909	10	1.0	
3	25	141.81	24.11	-0.0008	0.0008	18.897*	8	2.362	<0.01
4	25	215.24	34.31	-0.0002	0.0007	14.526*	6	2.421	<0.05

Mean slope of sessions 2 and 4 = 224.31. Square root of mean variance of slope for sessions 2 and 4 (S.E.) = 29.46

n=5

1	25	106.66	19.78	+0.0029	0.0012	34.702	14	2.479	<0.01
2	25	114.47	29.86	+0.0006	0.0012	30.580	10	3.058	<0.01
3	25	155.42	17.75	+0.0076	0.0008	19.756	14	1.0	
4	25	186.30	21.58	+0.0033	0.0005	17.918	10	1.0	

Mean slope of sessions 3 and 4 = 170.86. Square root of mean variance of slope for sessions 3 and 4 (S.E.) = 19.76

Observer G.S.

	SL	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	50	150	160.81	7.66	-0.0037	0.0002	11.742	13	1.0	
	40	150	208.71	9.00	-0.0021	0.0002	14.940*	9	1.0	
	30	150	290.96	11.91	-0.0014	0.0001	19.210	12	1.0	
	20	150	246.34	15.28	-0.0019	0.0002	19.805	9	2.201	<0.025
	15	150	208.80	9.04	-0.0023	0.0052	19.225	13	1.0	
n=5	50	150	153.31	10.04	+0.0042	0.0003	21.409	12	1.784	<0.025
	40	150	208.88	9.03	+0.0025	0.0002	16.836	13	1.0	
	30	150	242.22	14.40	+0.0022	0.0002	26.030	13	2.002	<0.01
	20	150	221.06	9.72	+0.0031	0.0002	8.236	13	1.0	
	15	150	198.73	9.08	+0.0038	0.0002	6.027	13	1.0	

Table 6.4. Probit analysis of observed psychometric functions from Experiment 6 - n is the lower harmonic number of the test signal. Observer A.F.									
	SL	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf
n=4	50	100	299.67	13.03	+0.0002	0.0001	9.328*	8	1.0
	40	100	349.25	15.29	+0.0001	0.0001	8.202	14	1.0
	30	100	353.38	15.62	+0.0000	0.0001	11.811	14	1.0
	20	100	306.56	13.32	+0.0001	0.0001	15.987	14	1.0
	15	100	259.62	11.78	+0.0000	0.0001	6.907	14	1.0
Observer G.S.									
	SL	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf
n=4	50	150	189.51	6.87	-0.0002	0.0002	14.009	14	1.0
	40	150	295.07	10.41	-0.0001	0.0001	22.546	14	1.0
	30	100	324.88	13.97	-0.0002	0.0001	20.457	14	1.0
	20	150	280.48	10.06	-0.0003	0.0002	20.627	14	1.0
	15	150	252.82	13.65	-0.0008	0.0002	22.744*	10	2.274

p<0.02:

The psychometric functions from Experiment 4 for the reference signal were well-fitted by a cumulative normal ogive. Since the two signals presented in the two observation intervals differed only in frequency, the variances $\sigma^2(x'_1)$ and $\sigma^2(x'_2)$ can be assumed to be equal, and the standard deviation of the internal representation of this signal was thus evaluated as the reciprocal of the slope of the fitted function divided by $\sqrt{2}$. The estimated standard deviations were 209.4 x 0.001813 Hz (0.3796 Hz) for observer A.F. and 209.4 x 0.002072 Hz (0.4339 Hz) for observer G.S.

The psychometric functions observed for Experiment 5 were analysed separately for each of the two comparison signals. For observer G.S., three of the ten observed functions were found to be significantly different from cumulative normal ogives with binomial variability with a criterion of p=0.050. These three functions were those for the comparison signal containing the 5th and 6th harmonics at 50 and 30 dB SL and that for the signal containing the 4th and 5th harmonics at 20 dB SL In the case of observer A.F., six of the ten observed functions were

significantly different from a cumulative normal ogive, these being the functions for each comparison signal at 20 dB SL,

the function for the comparison composed of the 4th and 5th harmonics at 15, 30 and 40 dB SL, and that composed of the 5th and 6th harmonics at 50 dB SL. However, with the exception of the functions from A.F. at 20 dB SL, no systematic deviations from the fitted functions are evident, and no importance is attached to these deviations. The performance of observer A.F. at 20 dB S.L. showed considerable between-session variability (see table 6.3), which was not evident at other signal levels. Both the slope and intercept changed from session to session, and the pooling of performance over sessions resulted in a comparatively large estimated standard deviation with a large associated standard error. In order to obtain a more realistic estimated standard deviation, the estimate for this condition was derived from the mean of the estimated slopes for the two sessions showing the largest slopes, and confidence limits were derived from the mean of the variances of these estimated slopes. Of the observed functions from Experiment 6, all but that for observer G.S. at 15 dB SL were well fitted by a cumulative normal ogive.

The Estimates of $\sigma(f'_0)/f_0$.

The parameter $\sigma(f'_0)/f_0$ was estimated from the data of Experiment 5 as $\sqrt{[\sigma^2(\underline{x}'_2 - \underline{x}'_1) - \sigma^2(\underline{x}'_1)]}$, where $\sigma(\underline{x}'_2 - \underline{x}'_1)$ was taken as the reciprocal of the slope of the theoretical function, and $\sigma(\underline{x}'_1)$ as the relative standard deviation of the internal representation of the reference signal

as estimated from the functions from Experiment 4. In the case of Experiment 6, $\sigma(f'_0)/f_0$ was estimated as $\sigma(x'_2 - x'_1)/\sqrt{2}$. The estimates of $\sigma(f'_0)/f_0$ as a function of the presentation level of the signal are shown for each observer in figure 6.7, together with 95% confidence limits for each estimate. In no case is the estimated $\sigma(f'_0)/f_0$ a monotonic function of the presentation level. This result is comparable with the results of Experiment 3, in which $\sigma(f'_j)/f_j$ was estimated for each component of a two-component complex signal under similar presentation conditions. As was pointed out above, some part of this non-monotonicity is likely to be due to the confounding of signal-to-noise ratio with the presentation level of the signal, but since the signal-to-noise ratio remained constant at about 40 dB for signal levels of 30 dB and above, this factor is not sufficient to explain the increase in the estimates of $\sigma(f'_0)/f_0$ at the higher signal levels. The presumed explanation of this increase is the same as that given for the comparable increase in the estimates of $\sigma(f'_j)/f_j$, that is, that the upward spread of masking from the low-pass noise has an effect whose magnitude increases more rapidly than the noise level for noise spectral-densities greater than about 30 dB SPL (Weber, 1977).

Comparison of $\sigma(f'_0)/f_0$ and $\sigma(f'_j)/f_j$.

The predicted relations between $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ derived in Chapter 2 from the spectral-pattern recognition and periodicity theories are shown in figure 6.8 together with the estimates of $\sigma(f'_0)/f_0$ for the signal where

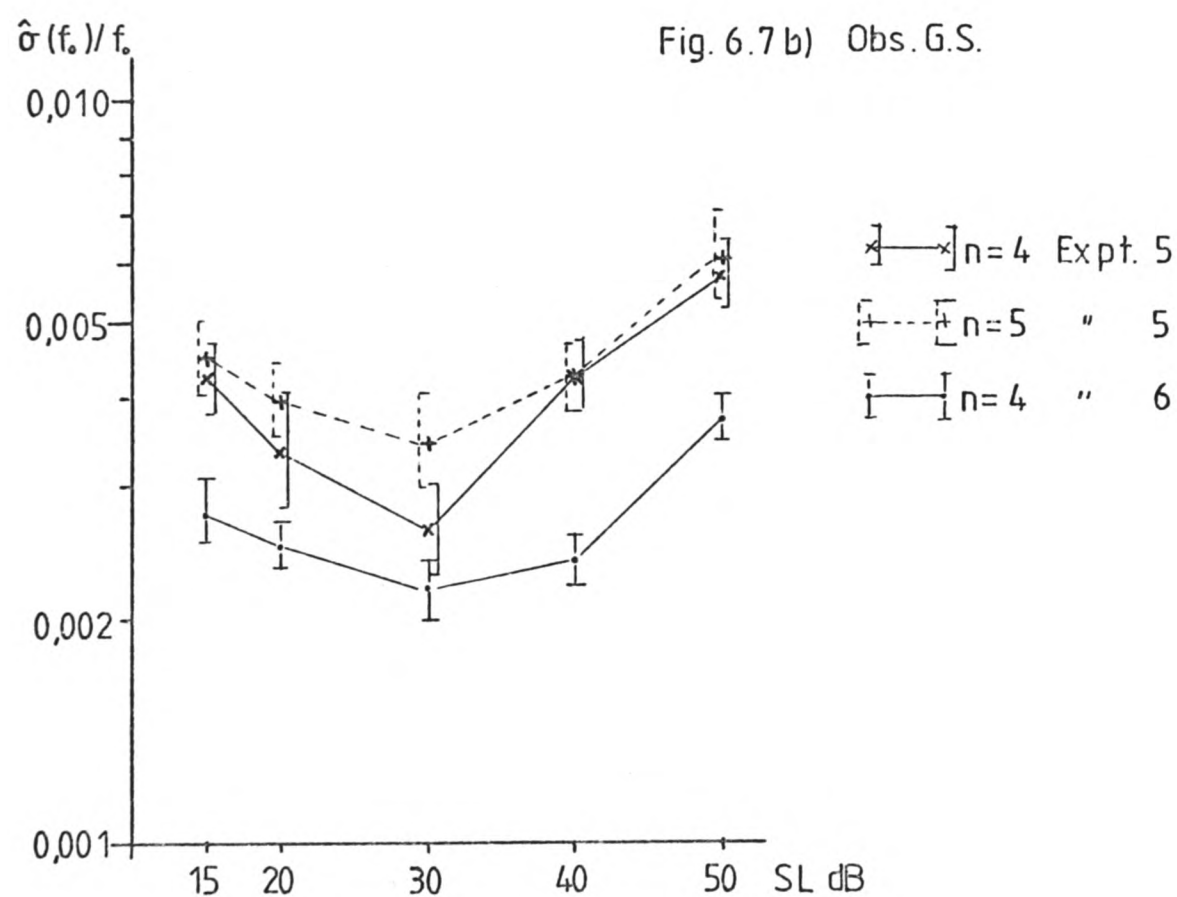
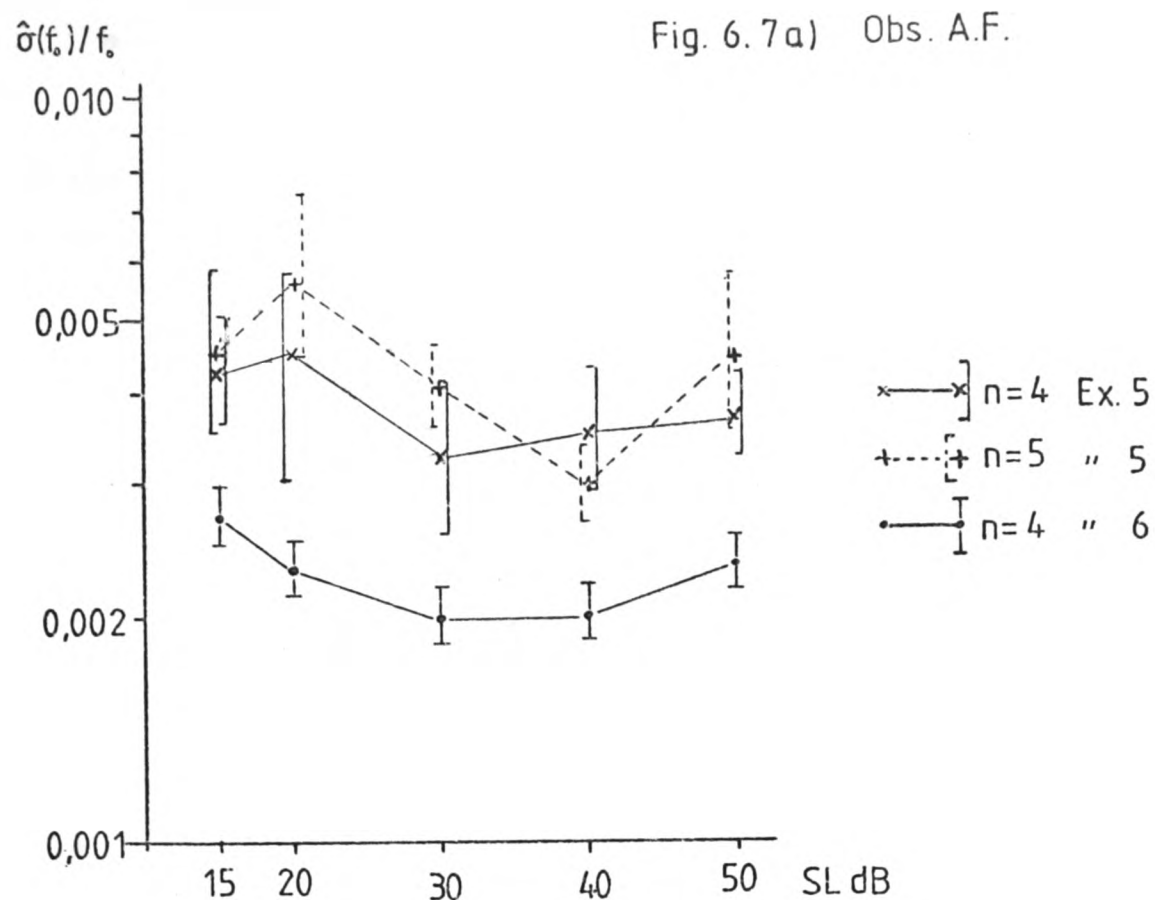


Figure 6.7. Estimates of $\sigma(f'_0)/f_0$ as a function of Sensation Level for two observers. The data are derived from Experiments 5 and 6. The legend n indicates the lower harmonic number of the two-component signal. Panel a), observer A.F., panel b), observer G.S. The error bars represent 95% confidence limits.

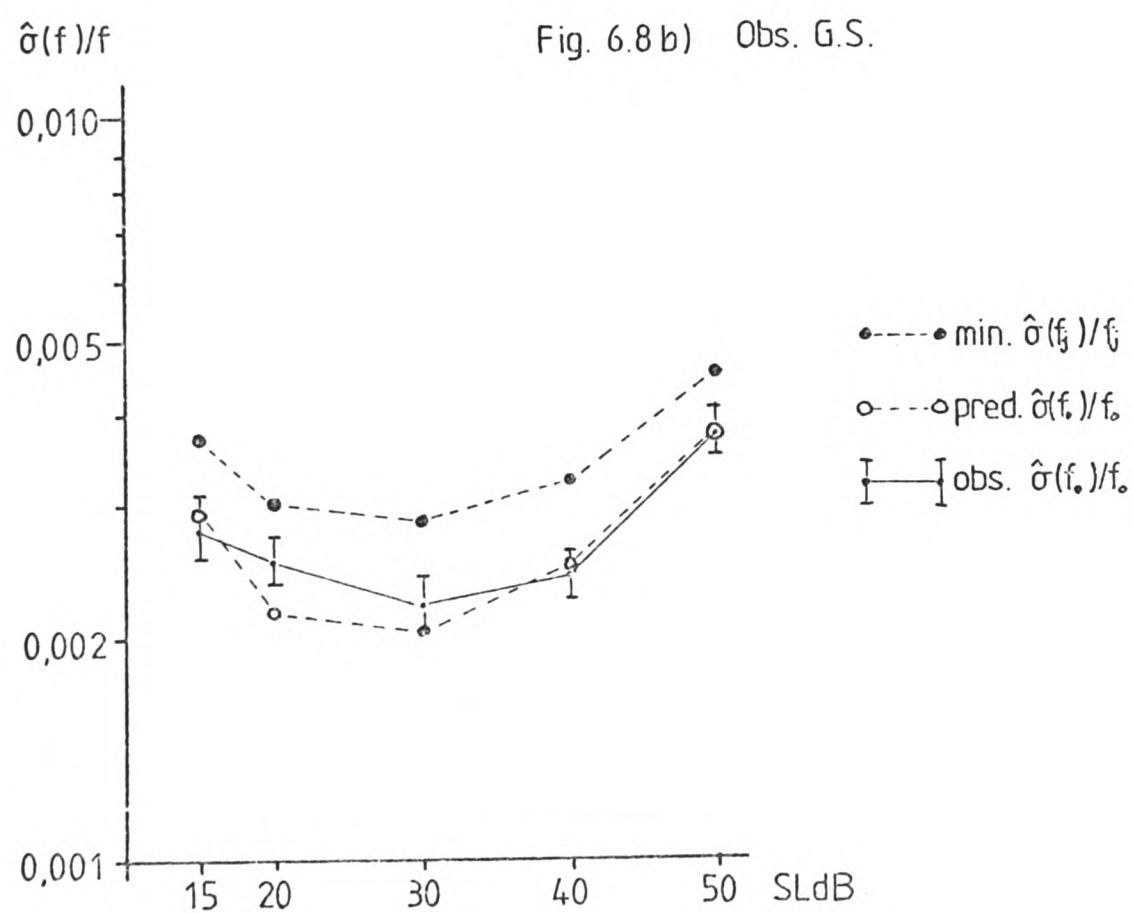
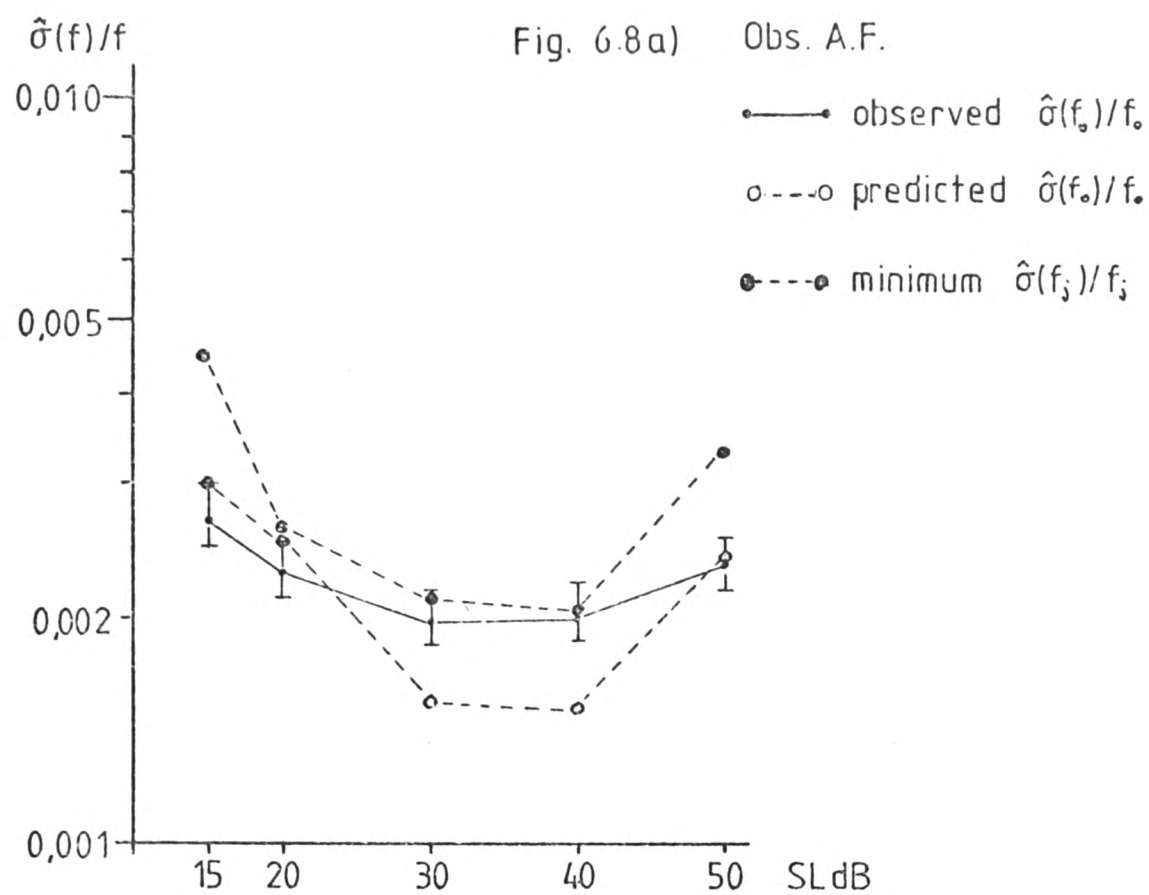


Figure 6.8. The estimates of $\hat{\sigma}(f'_0)/f_0$ from Experiment 6, together with the predictions of the spectral-pattern models based upon the estimates of $\hat{\sigma}(f'_j)/f_j$ from Experiment 3. The minimum value predicted by the periodicity theory is shown as the minimum of $\hat{\sigma}(f'_j)/f_j$. Panel a), observer A.F., panel b), observer G.S.

$\underline{n=4}$ from Experiments 5 and 6.

The relative DLs from Experiment 6 are consistently smaller than the smaller value of $\underline{\sigma(f'_j)/f_j}$ from Experiment 3 at the same presentation levels, and it can thus be concluded that performance in Experiment 6 was due to the use of information from both components rather than the comparison of just one of the component frequencies. Therefore, the DLs measured in Experiment 6 would seem to be due to a joint perception of both of the components of the signal.

The estimates of $\underline{\sigma(f'_0)/f_0}$ from Experiment 5 seem to show the same dependence upon the presentation conditions, as those from Experiment 6, but are somewhat larger than the values predicted by the spectral-pattern recognition theories, and also exceed the minimum values predicted from the periodicity theory. If the observers' performance in Experiment 6 was determined not by the residue pitch of the signal, but simply by a simultaneous comparison of the component frequencies, then the estimates of $\underline{\sigma(f'_0)/f_0}$ from Experiment 5 may indicate that here, $\underline{\sigma(f'_0)/f_0}$ is determined by the variability of periodicity information rather than the variability of independent estimates of the component frequencies.

The discrepancy between estimates of $\underline{\sigma(f'_0)/f_0}$ from the two experimental methods.

If stimulus uncertainty contributes additional noise to the estimates of the two component frequencies, then the

spectral-pattern recognition models would predict a difference between the estimates of $\underline{\sigma(f'_0)}/f_0$ from these two tasks. The task for Experiment 5 requires the monitoring of three frequency selective channels during the presentation of the two-component signal, so that three channels contribute noise to $\underline{\sigma(f'_0)}/f_0$. The signal for Experiment 6 always contained components in the same two frequency regions, so that only these two channels need contribute noise. This uncertainty difference leads to a predicted difference of a factor of $\sqrt{3/2}$, or about 1.22. The average factors by which the estimated values of $\underline{\sigma(f'_0)}/f_0$ from Experiment 5 exceed those from Experiment 6, are, for observer A.F., 1.71 for the complex composed of the 4th and 5th harmonics, and 2.09 for the complex composed of the 5th and 6th harmonics. The corresponding values for observer G.S. are 1.51 and 1.66 respectively. At 20 and 30 dB SL, with the signal from Experiment 5 being that containing the 4th and 5th harmonics, the factors for observer G.S. are 1.38 and 1.25, which are close to the theoretical factor of 1.22.

The differences between the estimates of $\underline{\sigma(f'_0)}/f_0$ from these two different tasks could be explained by an additional internal variance, since the differences between the squares of these estimates are approximately constant. The square root of the mean of the differences between the squares of the estimates at the same SLs are 0.003515 for observer A.F. and 0.003483 for observer G.S. The estimated standard deviations obtained here do not differ greatly as a function of SL, and for this reason it is difficult to evaluate whether the two sets of estimates are related by an additive factor, by a multiplicative factor,

or by some combination of the two. This issue is addressed below (see Chapter 9), after a number of other experiments intended to estimate $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ have been described.

A third possible explanation of this discrepancy is suggested by the comparison of the masking effects of gated and continuous noise described by Weir, Green, Hafter, and Burkhardt (1977). This study showed that a gated noise masker tends to be more effective than a continuous noise masker of the same power. Thresholds in gated noise were higher than in continuous noise, particularly for short (10 msec) duration signals and at frequencies of 1000 Hz and higher. The threshold differences were, however, observed with a mask gated on simultaneously with, or only 20 msec before, the target, and the observed threshold differences were greater for simultaneous onset of target and mask. In this study, the noise was gated on 100 msec before the signal onset, and it is considered unlikely that an effect of this sort is of great importance here.

General Comments.

These results confirm the suggestion introduced above that the variability of the estimates of component frequency upon which the estimates residue pitch appears to be based is dependent upon factors other than frequency. One point raised earlier was that the presence of the 4th harmonic, which is related in pitch to the fundamental frequency by a difference of two octaves, might itself be a sufficient cue for pitch

discrimination in a tasks such as those above. The results of Experiment 5 allow a comparison of pitch discrimination performance for two complex signals, of which only one contains the 4th harmonic. A comparison of the estimates of $\sigma(f'_0)/f_0$ from these two complexes from Experiment 5 may be made from figure 6.7. While these estimates do tend to be slightly smaller for the complex containing the 4th harmonic, differences of this order do not suggest that the pitch of this component itself is the primary cue for the observer's judgements.

Summary.

The results of Experiments 3 and 6 taken in combination appear to be quantitatively consistent with the predictions of the spectral-pattern recognition theories, but the results of Experiment 5 raise some doubt as to whether the pitch discrimination measured in Experiment 6 is due to the residue pitch of the signal. The estimates of $\sigma(f'_0)/f_0$ from Experiment 5 are somewhat larger than would be predicted by these theories if the results of Experiment 3 do indeed represent estimates of $\sigma(f'_j)/f_j$ as seen by the pitch processor. An additional reason for a cautious interpretation of these results is that the range over which the parameter $\sigma(f'_j)/f_j$ could be manipulated by the range of signal levels employed here is rather small. For this reason, the experiments described in the following chapters employed alternative independent variables in an attempt to increase this range.

CHAPTER 7

Abstract.

This chapter contains a description of five experiments intended to estimate the effects of the signal energy to noise spectral-density ratio ($E-N_o$) upon the standard deviation of the internal representation of the component frequencies and residue pitch of a two-component harmonic signal. The results of the experiments described in Chapters 4 - 6 show only a small effect of presentation level upon frequency discrimination, which may have been due to the approximately constant signal-to-noise ratio at the different presentation levels. The use of an external noise has the advantage that the contribution of the external noise may be considerably greater than the effect of any internal noise which may be present, so that the effective level of the internal noise may be estimated from the form of the dependence of discrimination upon the external noise level. The use of $E-N_o$ as an independent variable also allows a test of the predictions of the models of frequency discrimination discussed in Chapter 3 for the case of residue

pitch.

Experiment 7, the first of the experiments described here, is similar to Experiment 3, and consisted of a 2AFC frequency discrimination task for each component of a two-component harmonic complex, the test signal being a sinusoidal tone-burst presented together with a continuous sinusoid, a low-pass noise, and a white noise of variable spectral density. This experiment resulted in series of estimates of $\sigma(f_j')/f_j$ for each component of the signal as a function of $E-N_0$. Two of the experiments, Experiments 8 and 9, were essentially similar to Experiment 5. The observer's task was to compare the pitches of a five-component reference signal and a two-component comparison signal, the two signals having no harmonics in common. Experiment 10 was essentially similar to Experiment 6; here the observer's task was to compare the pitches of two 2-component signals constructed from the 4th and 5th harmonics of two similar fundamental frequencies. These three experiments resulted in estimates of $\sigma(f_0')/f_0$ as a function of $E-N_0$.

The estimates of $\sigma(f_0')/f_0$ derived from Experiment 10 were consistent with the predictions of the spectral-pattern recognition theories based upon the estimates of $\sigma(f_j')/f_j$ from Experiment 7. The estimates of $\sigma(f_0')/f_0$ derived from Experiments 5 and 6 described above differed, as do the estimates derived here from Experiments 8 and 10, and this

difference was ascribed to the effects of stimulus uncertainty. This explanation is tested here. Experiments 8 and 9 differed in that in the former experiment, there were two possible comparison signals with different harmonic numbers, while in the latter experiment, only one comparison signal was presented. Thus, the stimulus uncertainty inherent in Experiments 5 and 8 is not present in Experiment 9. The effect of this uncertainty upon the estimates of $\sigma(f'_0)/f_0$ is that the estimates of $\sigma(f_0)/f_0$ from Experiment 8 exceed those from Experiment 9 by an average factor of 1.490, which is approximately consistent with the factor of 1.22 predicted from the uncertainty difference.

If the only cause of the difference between the estimates of $\sigma(f'_0)/f_0$ from Experiments 5 and 6 was a comparable uncertainty difference, then the estimates of $\sigma(f_0)/f_0$ from Experiments 9 and 10 should be equal. The obtained estimates differ, those from Experiment 9 being on average 1.556 larger than those from Experiment 10. It thus appears that an observer's ability to make pitch comparisons is better when the pitches are coded in the same frequency selective channels. This result has two possible explanations: Either the comparison of pitches derived from components of different frequencies is carried out by the use of periodicity information, or some internal noise process degrades temporally or spectrally coded estimates of component frequency. The latter explanation is consistent with Goldstein's (1973)

assumption that an internal noise process limits the precision of coding of residue pitch.

The fourth experiment, Experiment 11, was intended to test the hypothesis that between-trial comparisons are at least partially responsible for the discrimination performance of the two observers in the experiments employing a reference signal. If this hypothesis is true, there should be significant differences in frequency discrimination performance between a task where the absolute value of the fundamental frequency difference Δf is constant within a block of trials, and a task where a number of different values of Δf occur randomly within a block of trials. Experiments 8 and 11 differed in just this respect, and a comparison of the results of these experiments shows that between-trial comparisons are not an important determinant of the discrimination performance measured here, since the random presentation of different values of Δf within a block of trials has only slight effects upon the observed performance.

The results of the experiments indicate that an Energy Detector model of frequency discrimination is sufficient to account for the dependence on $E-N_0$, at least where $E-N_0$ is between about 20 to 35 dB.

Introductory Comments.

If the residue pitch of a two-component complex signal is derived from independent internal representations of the frequencies of the two components, the models of frequency discrimination described in Chapter 3 can predict the relationship between $E-N_0$ and the standard deviation of the internal representation of the component frequencies, and thus, from the spectral-pattern recognition models, the dependence of the standard deviation of the internal representation of the residue pitch of the signal upon $E-N_0$ can also be predicted. According to the Energy Detector model, if the signal level is constant and the noise level is varied, this standard deviation (see equation 3.19) should be a function of $1/(N_0 + k \sqrt{N_0})$, where N_0 represents the noise spectral-density, the average noise power in a 1 Hz band, in the region of the signal components, and k is a constant dependent upon the bandwidth of the filters at whose outputs energy is measured. The data of Henning (1967b) for single sinusoids are generally consistent with this relationship where $E-N_0$ is less than 35 - 40 dB. At higher values of $E-N_0$, frequency discrimination no longer improves with $E-N_0$, but remains approximately constant; the Energy Detector model can, however, account for this result if it is assumed that an internal noise process with a level equivalent to a spectral-density of 35 - 40 dB ^{below the signal power} limits performance.

The model described by Siebert (1970), which is based upon the use of the timing information in the firing pattern of the auditory nerves, predicts that this standard deviation should be a function of $1/\sqrt{\ln(A)}$, where A is the amplitude of the signal components in units of threshold amplitude, where that threshold will here be determined by noise. For most observers, the ratio of signal energy to noise spectral-density at threshold is approximately 16 dB (Green, 1976). For values of $E-N_0$ of between 20 and 40 dB, the forms of the effects of $E-N_0$ upon frequency discrimination predicted by these two models are rather similar, and it is unlikely that data of sufficient precision to discriminate between the predicted forms can be collected. The actual values predicted by Siebert's model are some two orders of magnitude smaller than those generally obtained however, and it is clear that this model is not appropriate as a description of human performance. The temporal model described by Luce and Green (1974) predicts the same form for the amplitude dependence of the DL as Siebert's model, and since Luce and Green assume a less than optimally efficient decision process, the DLs predicted by their model are similar to those found empirically. Goldstein and Srulovicz's (1977) temporal model cannot yield direct predictions for the effects of signal and noise levels, since these factors are assumed to have some unspecified effect upon the number of active auditory fibres, their firing rate, and the degree to which their activity is synchronised to the signal.

These experiments are thus unlikely to allow the rejection of either the Energy Detector model or of the temporal models described by Luce and Green (1974) and by Goldstein and Srulovicz (1977). However, when considered together with the effect of signal duration upon pitch discrimination for the residue pitch of a complex signal presented in noise, these results will allow a critical test of the Energy Detector model, which predicts that the DL is primarily determined by the difference between the signal and noise energies. Since signal energy is inversely proportional to signal duration, provided that the integration time remains constant, the noise energy is independent of signal duration; thus, the effect of varying signal duration can be considered as a variation in the energy difference between signal and noise, and for any signal duration less than the integration time and greater than the time constant of the assumed filter, the DL should be constant for a given difference between the signal and noise energies. Since the temporal models of frequency discrimination derive their predictions for the effects of signal duration and signal-to-noise ratio from independent terms, such an equivalence does not follow from these models. Two experiments investigating the effect of signal duration upon frequency discrimination for residue pitch are described in the following chapter.

The first four experiments described here are essentially

similar in design to Experiments 3, 5 and 6, which were described in Chapters 5 and 6. The methods, and the reasons for the choice of these methods, are identical.

Experiment 7.

Observers.

Only one of the two observers, A.F., was available for this experiment.

Stimuli

This experiment was similar to Experiment 3, which was described in Chapter 5. The test signal was a sinusoidal tone-burst at the frequency $\underline{n}(209.4 + \underline{\Delta f}/2)$ or $\underline{n}(209.4 - \underline{\Delta f}/2)$ Hz, where \underline{n} was equal to 4 or 5. The background to the presentation of the tone-bursts consisted of a continuous sinusoid at $\underline{m}(209.4)$ Hz, where $\underline{m}=5$ if $\underline{n}=4$, and $\underline{m}=4$ if $\underline{n}=5$, a low-pass filtered noise, and a white noise whose spectral density was 20, 25, 30, 35, or 40 dB below the level of the test signals. The test signals and the continuous sinusoid were of constant and equal level, such that the signal at 5×209.4 Hz was about 30 dB above the observer's threshold in quiet. The time-course of a single trial is illustrated in figure 7.1.

The low-pass filtered noise was filtered at a cut-off frequency of 613 Hz with a slope of 72 dB/octave. The spectral-density of the noise at 618 Hz, the frequency of the cubic-difference-tone $2f_1 - f_2$, was 10 dB below the level of the sinusoids at 837.6 and 1047 Hz, while at the latter two frequencies, the spectral-density of this noise was 40 dB below the level of these two components. The white noise was taken from the same source as the low-pass filtered noise. All level measurements were made electrically at the earphone input with a wave analyser (HP 3580a) set for a 1 Hz measuring bandwidth.

Procedure

Trials were presented in blocks of 50 with $|\Delta f|$ constant. Eight blocks, with values of $|\Delta f|/f$ from 0.03248 to 0.00101, comprised a test run, which lasted for about 25 minutes. In a given run, the value of $|\Delta f|$ either increased or decreased over blocks. Four test runs were given for each value of n and each white noise level, resulting in 200 observations at each value of $|\Delta f|$ within each of these treatments. The harmonic number n , the white noise level, and the order of change of $|\Delta f|$ over blocks were counterbalanced. One stimulus tape was prepared for each of the two test signals, and was replayed both forwards and backwards.

Experiment 8.Observers

Both of the observers A.F. and G.S. served in this experiment.

Stimuli

The first signal, the reference signal, presented on each trial was the signal presented in Experiment 4, except that the fundamental frequency of this signal was always equal to 209.4 Hz. Each component of this signal was presented at a level 40 dB above the observer's threshold at 1047 Hz in a silent background. During the presentation of the reference signal, no background noise was present. The second signal, the comparison signal, contained either the 4th and the 5th or the 5th and the 6th harmonics of $209.4 + \Delta f$ Hz. Each of the two components was presented at the same sound pressure level, this being 30 dB above the observer's threshold at 1047 Hz with a silent background. The noise background was introduced after the first 100 msec of the 200 msec inter-stimulus interval, and was removed again 100 msec after the offset of the comparison signal. The time-course of a trial is illustrated in figure 7.2.

Figure 7.3. The time-course of a single trial for Experiment 10.

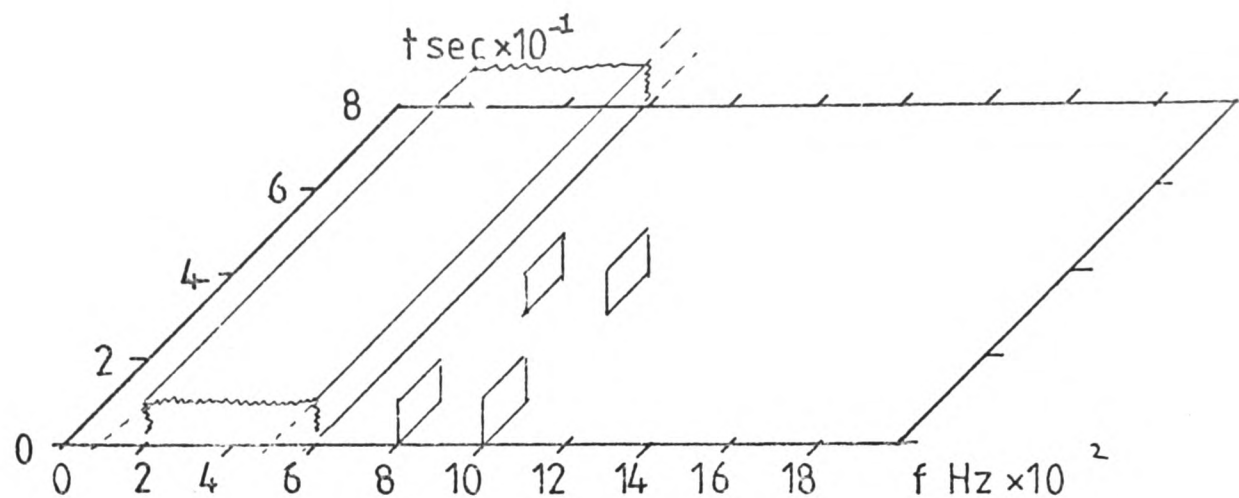


Figure 7.2. The time-course of a single trial for Experiment 8.

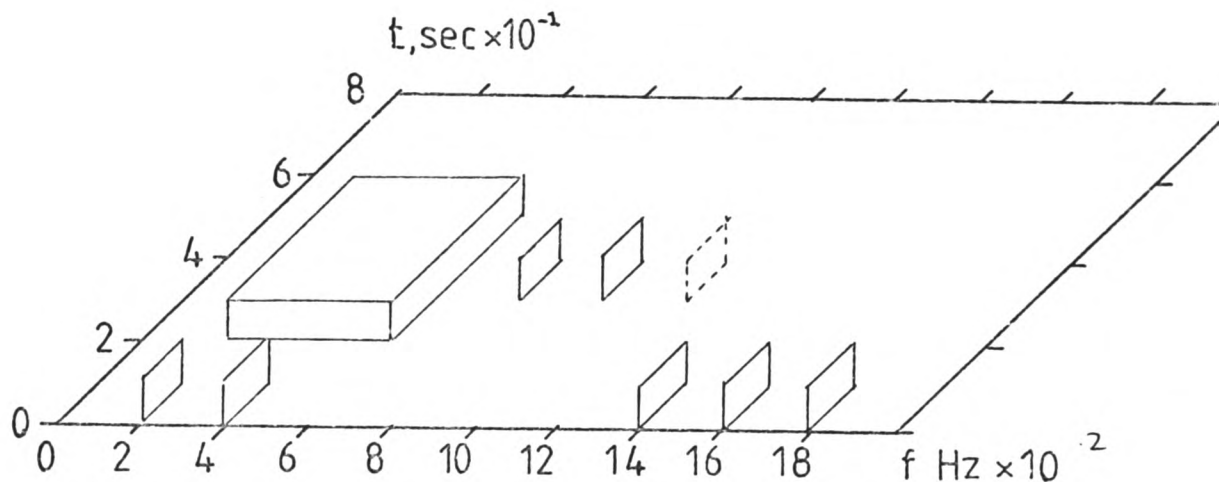
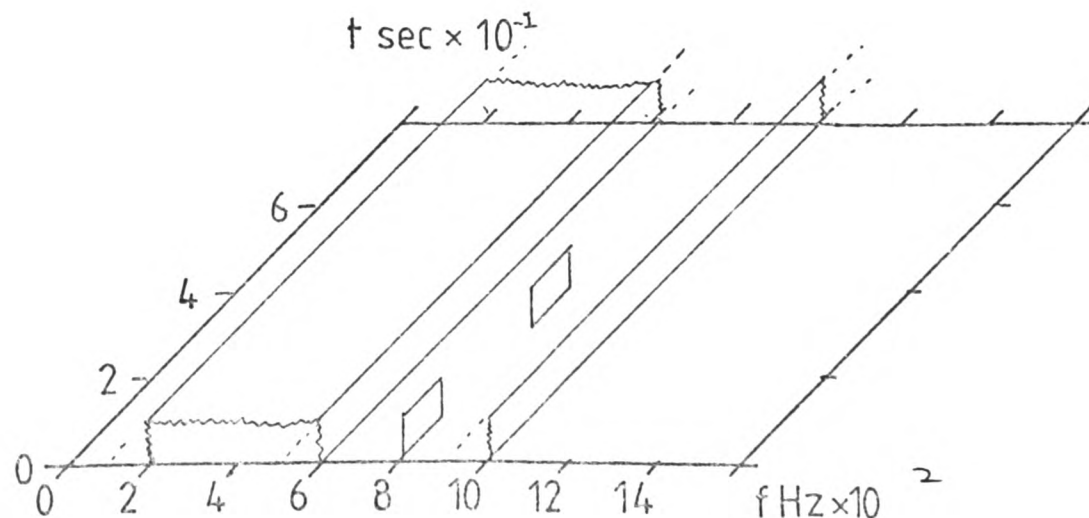


Figure 7.1. The time-course of a single trial for Experiment 7.



The noise consisted of the sum of a white noise whose spectral-density was varied between 20 and 40 dB relative to the components of the comparison signal, and a low-pass filtered noise intended to mask any distortion products lying below the components of the comparison signal. These two noise signals were both derived from a single white noise generator. The cut-off frequency for low-pass filtering was at 613 Hz, and the filter slope was 72 dB/octave. The spectral-density of the low-pass filtered noise at 613 Hz was 20 dB below the levels of the components of the comparison signal as measured electrically at the headphone input with a wave analyser (HP 3580A) set for a 1 Hz measuring bandwidth. The measured noise level was taken to be the mean, estimated by eye, of the readings of the analyser; the peak noise levels were also noted, these typically being about 5 dB above the mean reading.

Trials were presented in blocks of 100 with $|\Delta f|$ constant. The value of $|\Delta f|$ ranged from 209.4×0.03248 Hz to 209.4×0.00101 Hz. Within a block, the probability that the comparison signal had a higher fundamental frequency than the reference signal was 0.50, the probability that the comparison signal contained the 4th and 5th harmonics was 0.50, and these two events were independent. A series of eight blocks with a constant white noise level was presented in a test session of about one hour in length. Two stimulus tapes were prepared, with blocks of trials having the value of $|\Delta f|$ increasing and

decreasing. The order of test sessions at the different signal-to-noise ratios and orderings of $|\Delta f|$ was counterbalanced over the days of presentation. Each observer received six series of blocks at each signal-to-noise ratio, and thus received 300 trials for each of the two comparison signals at each value of $|\Delta f|$ at each of the signal-to-noise ratios.

Experiment 9.

This experiment was identical to Experiment 8 above, except that the lower harmonic number of the comparison signal was always equal to 4, and trials were presented in blocks of 50. Only observer A.F. was available when the experiment was run.

Experiment 10.

This experiment was similar to Experiment 6 above, except that the signal level was constant at 30 dB SL, and the noise background contained a white-noise, having a spectral-density varying from 20 to 40 dB relative to the level of the signal components, in addition to the low-pass filtered noise.

Observers

Both of the observers A.F. and G.S. served in this experiment.

Stimuli.

Both of the stimuli presented on each trial contained the 4th and 5th harmonics of $209.4 + \Delta f/2$ or $209.4 - \Delta f/2$ Hz. The two components of each signal were presented at the same sound pressure level, which was 30 dB above the observers threshold in quiet at 1047 Hz. These signals were presented against the same background of white and low-pass filtered noise as that presented in Experiment 8, except that this noise was on continuously. The structure of a single trial is illustrated in figure 7.3.

Each block contained 50 trials, and the probability that the second signal had the higher fundamental frequency was 0.50. The value of $|\Delta f|$ was varied from 209.4×0.01218 Hz to 209.4×0.00101 Hz. Each test session lasted about an hour, and comprised two series of eight blocks, the level of the white noise being constant within each series. The value of $|\Delta f|$ either increased or decreased over the blocks of a series, and the ordering over test sessions of blocks at different noise levels and with the $|\Delta f|$ increasing or decreasing was counterbalanced. Each observer received at least 200 trials at each value of $|\Delta f|$ of presentation level. A single stimulus tape was prepared, and was replayed in each of the two possible directions.

Experiment 11.Observers.

Both of the observers A.F. and G.S. served in this experiment.

Stimuli and Procedure.

This experiment was similar in design to Experiment 8, the signals and background noise being identical to those used in that experiment. This experiment differed in that the value of $|\Delta f|$ was not constant within a block of trials, but varied randomly between five possible values. These values, expressed as a proportion of the fundamental frequency, were equal to 0.03248, 0.01624, 0.00812, 0.00406 and 0.00203. Each test session comprised five blocks of 100 trials. Each block contained 25 trials at each of the five values of $|\Delta f|$; within these 25 trials, Δf was positive with a probability of 0.50, and independently of the sign of Δf , the lower harmonic number of the comparison signal was equal to 4 on half of the trials, and equal to 5 on the other half. Each observer received three such test sessions at values of $E-N_0$ between 20 and 40 dB, and thus received 150 trials for each of the two comparison signals at each value of $|\Delta f|$ and at each signal-to-noise-ratio.

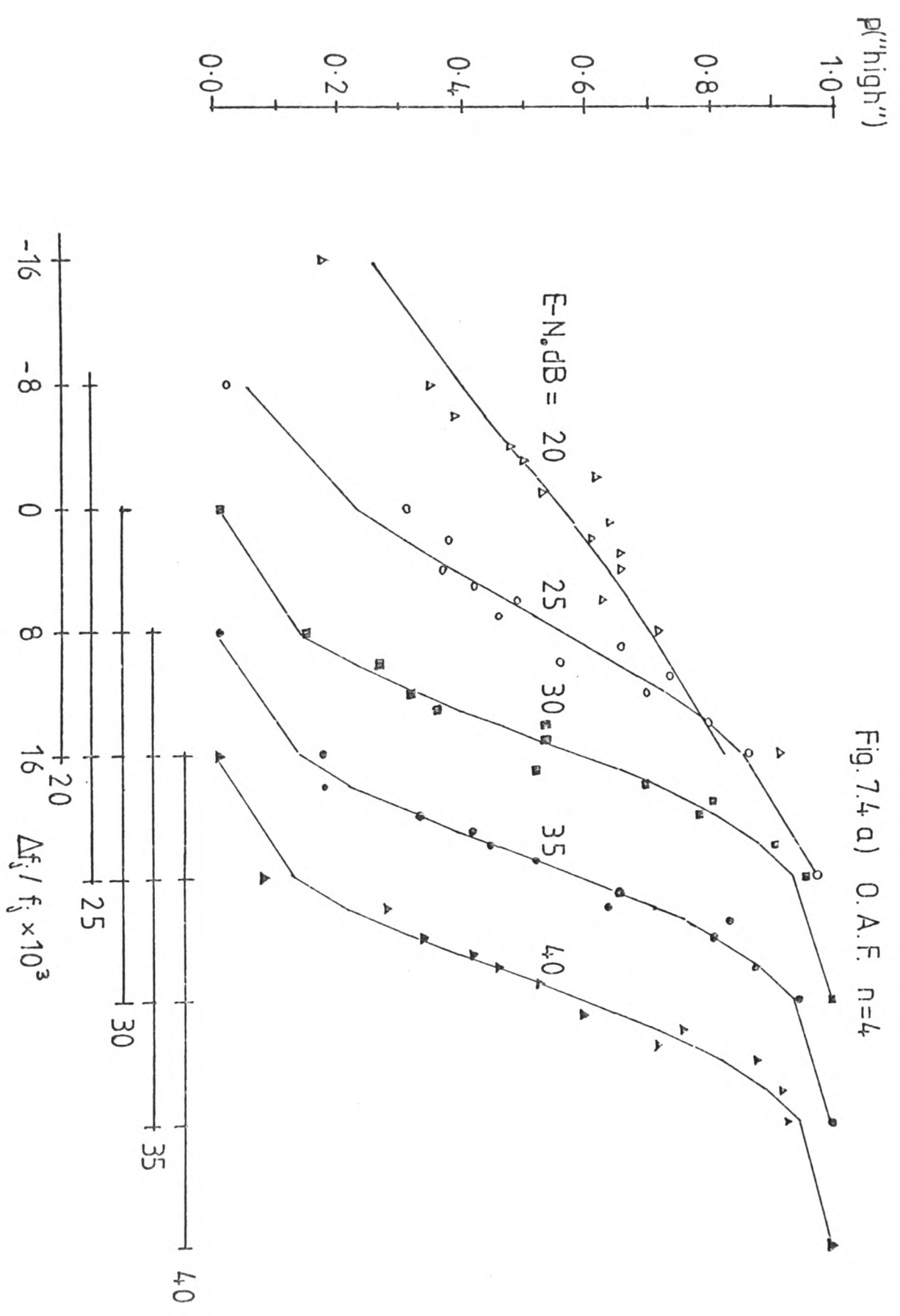
Results.

Psychometric Functions.

The observed psychometric functions were in each case subjected to a Probit analysis (Finney 1971), from which maximum-likelihood estimates of the slope and intercept of each function were obtained for performance as a z-score of proportions from 0 to 1, as a function of the relative frequency difference $\Delta f/f_0$ scaled linearly. The observed and theoretical functions appear as figures 7.4, 7.5, 7.6, 7.7 and 7.8. The parameters of the fitted functions are displayed in tables 7.1, 7.2, 7.3, 7.4 and 7.5.

Figure 7.4. Psychometric functions from Experiment 7; frequency discrimination as a function of signal-to-noise ratio for each of the 4th and 5th harmonics of 200 Hz. Panel a), harmonic number \underline{n} of the test signal is 4. Panel b), $\underline{n} = 5$. 100 trials per point, data from observer A.F. only. The solid lines are drawn between the best-fitting theoretical function.

Symbol	E-N ₀ dB
Δ	20
○	25
■	30
●	35
▲	40



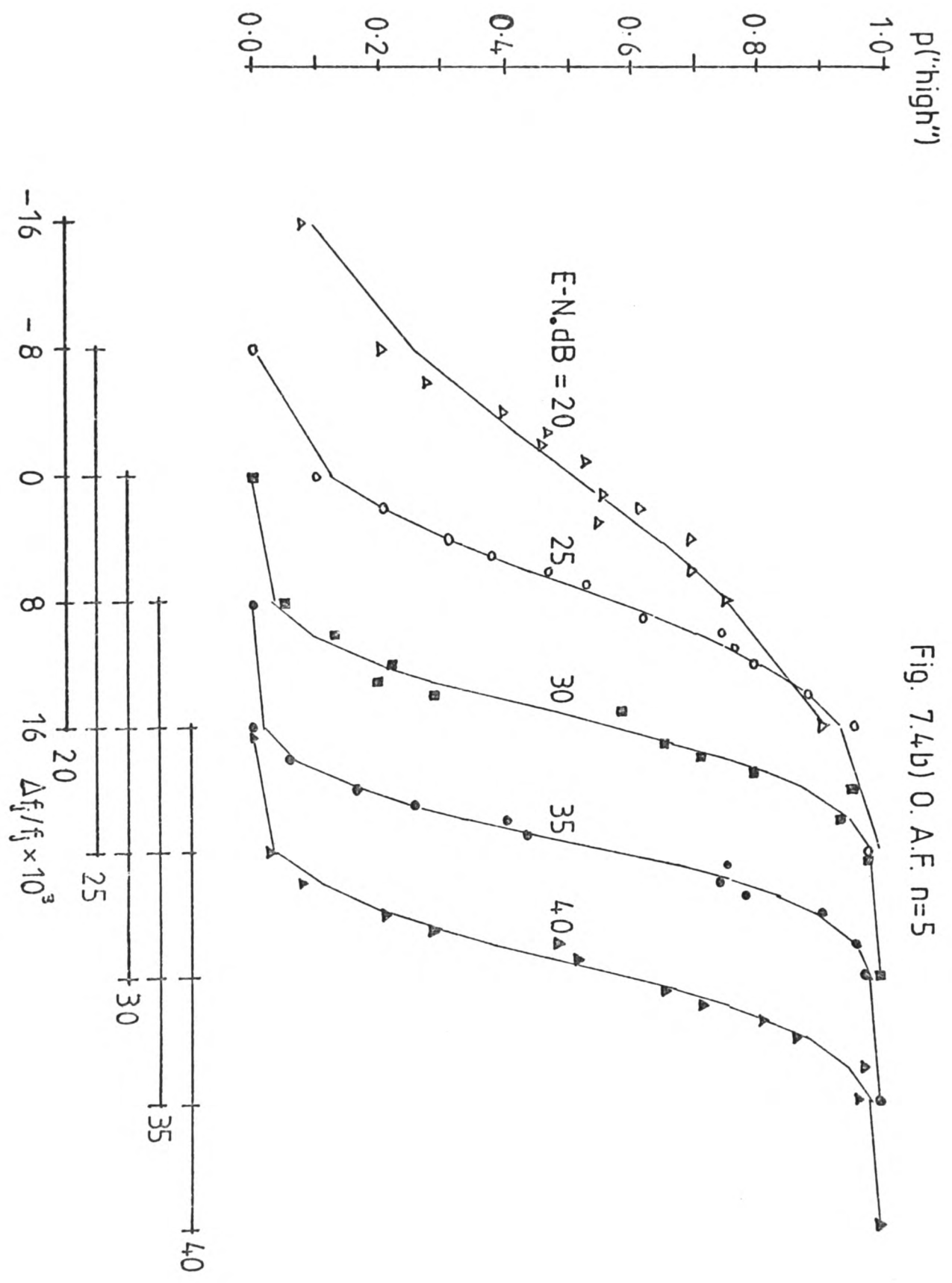


Figure 7.5. Psychometric functions from Experiment 8; pitch discrimination between the reference signal and the two-component comparison signal.

Panel a), observer A.F., $\underline{n} = 4$.

Panel b), observer A.F., $\underline{n} = 5$.

Panel c), observer G.S., $\underline{n} = 4$.

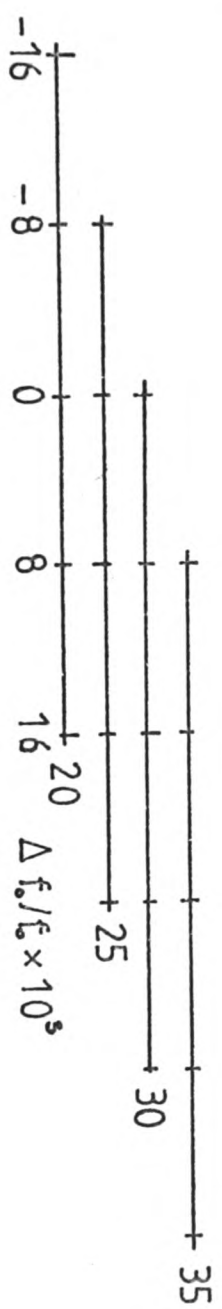
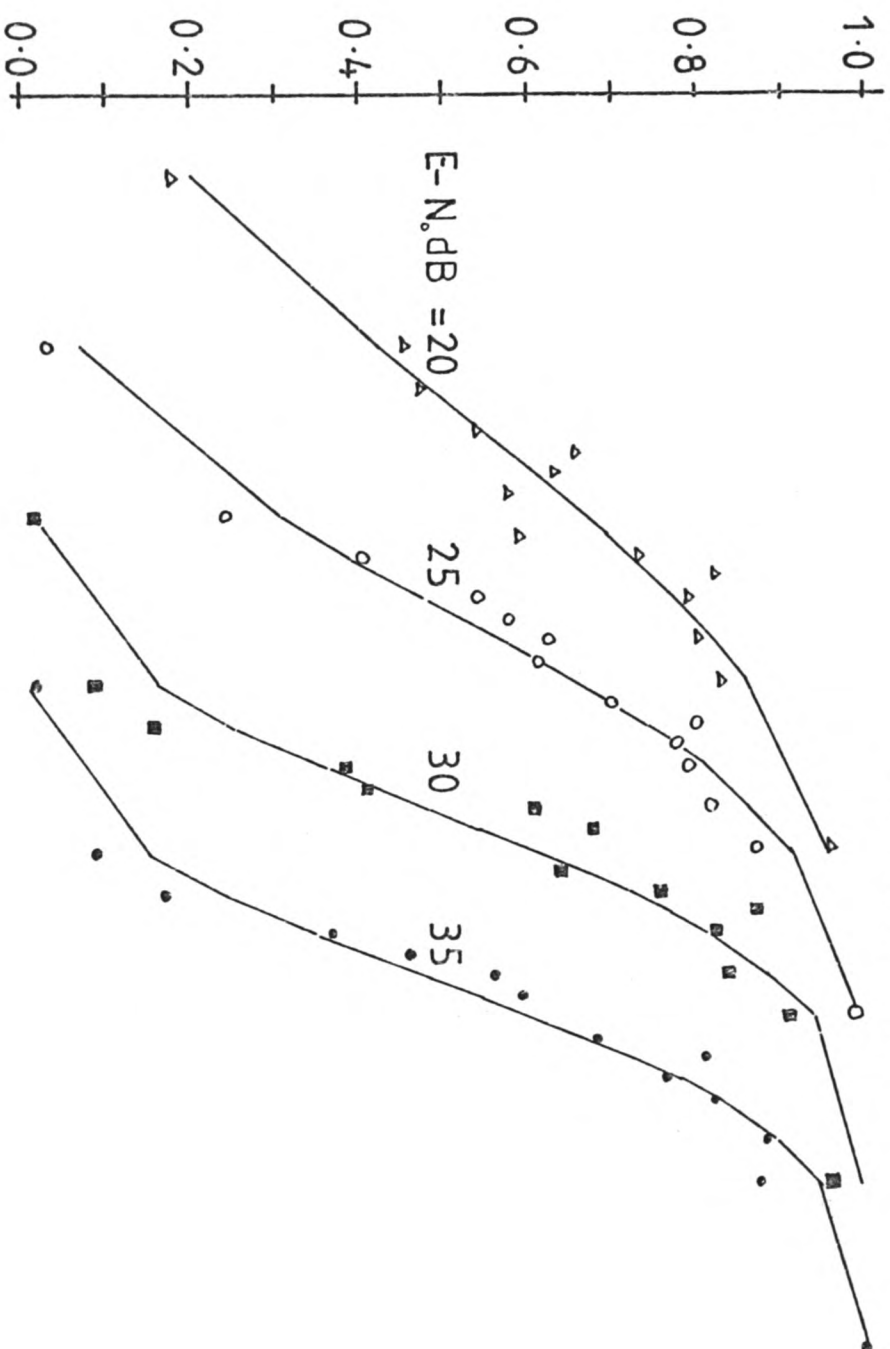
Panel d), observer G.S., $\underline{n} = 5$.

The legend \underline{n} represents the lower harmonic number of the comparison signal. The symbols have the same meaning as those in figure 7.4.

All points are based upon 150 trials except for those from observer G.S. at $E-N_0 = 20$ dB, where there are 75 trials per point.

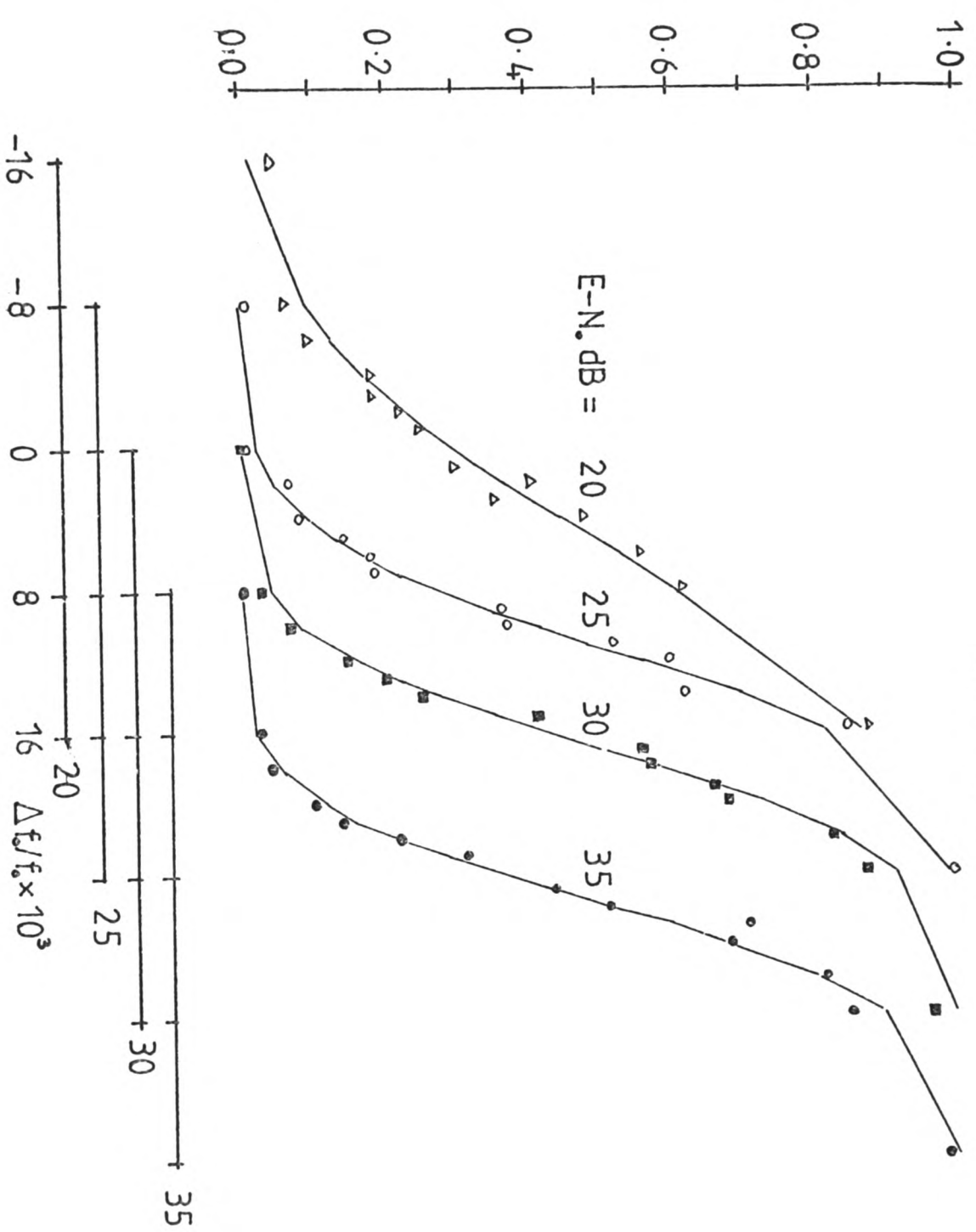
$P(\text{"high"})$

Fig. 7.5 a) 0. A.F. $n=4$



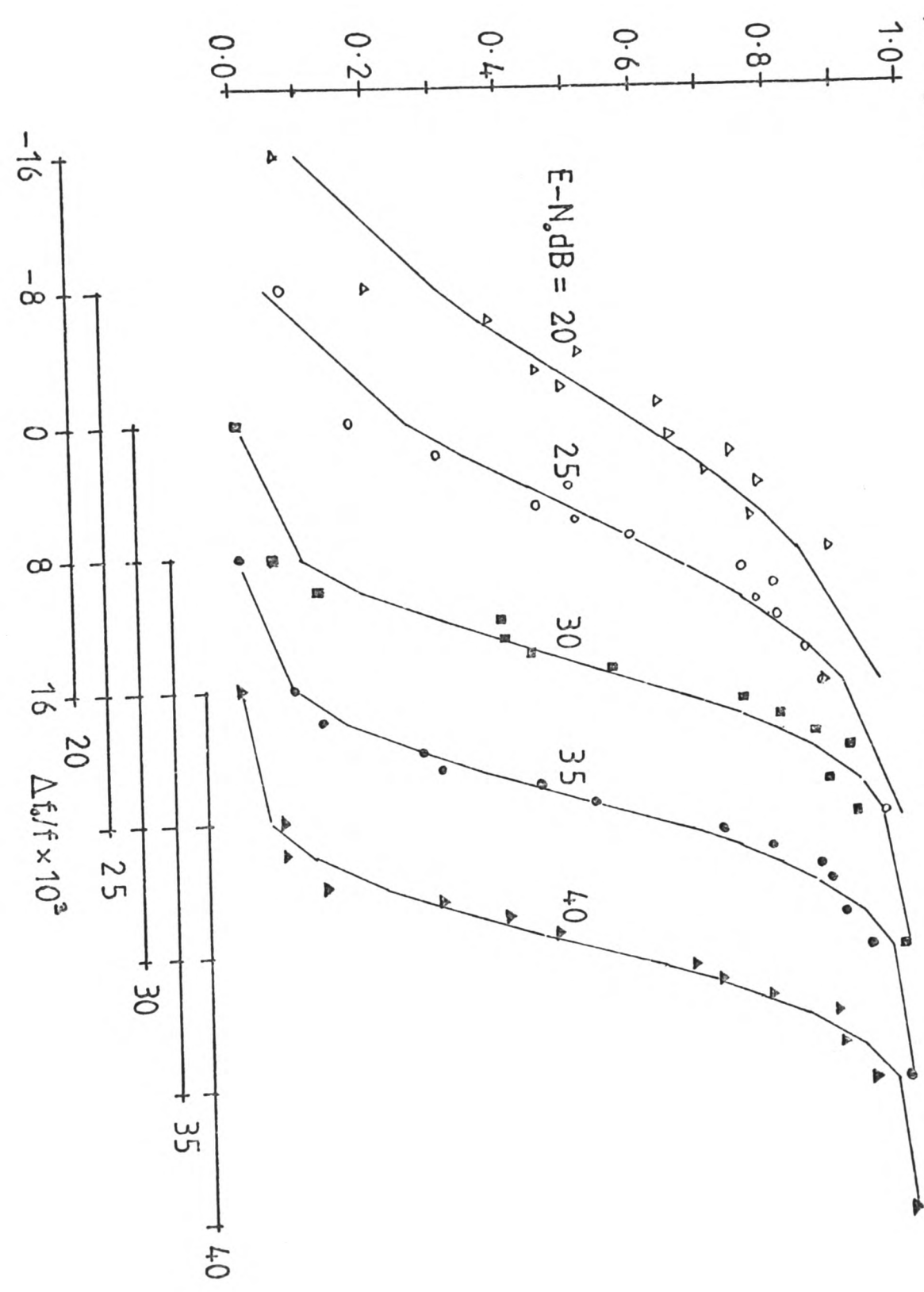
P ("high")

Fig. 7.5 b) 0. A.F. $n=5$



$P(\text{"high"})$

Fig. 7.5c) 0. G.S. $n=4$



P("high")

Fig. 7.5 d) 0. G.S. n=5

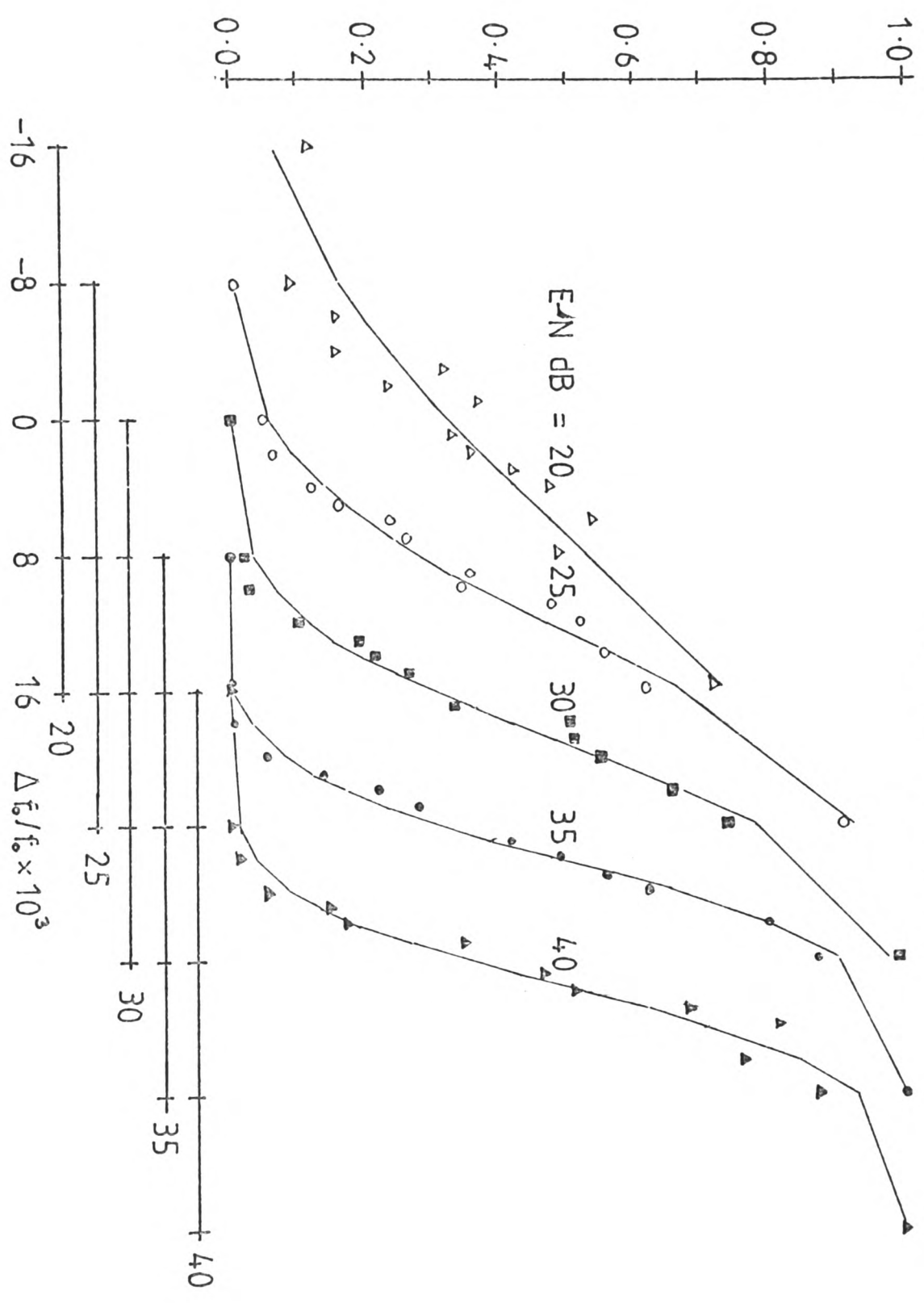


Figure 7.6. Psychometric functions from Experiment 9; pitch discrimination between the reference signal and a two-component comparison signal having $\bar{n} = 4$. Observer A.F. only, 100 trials per point. Symbols as figure 7.4.

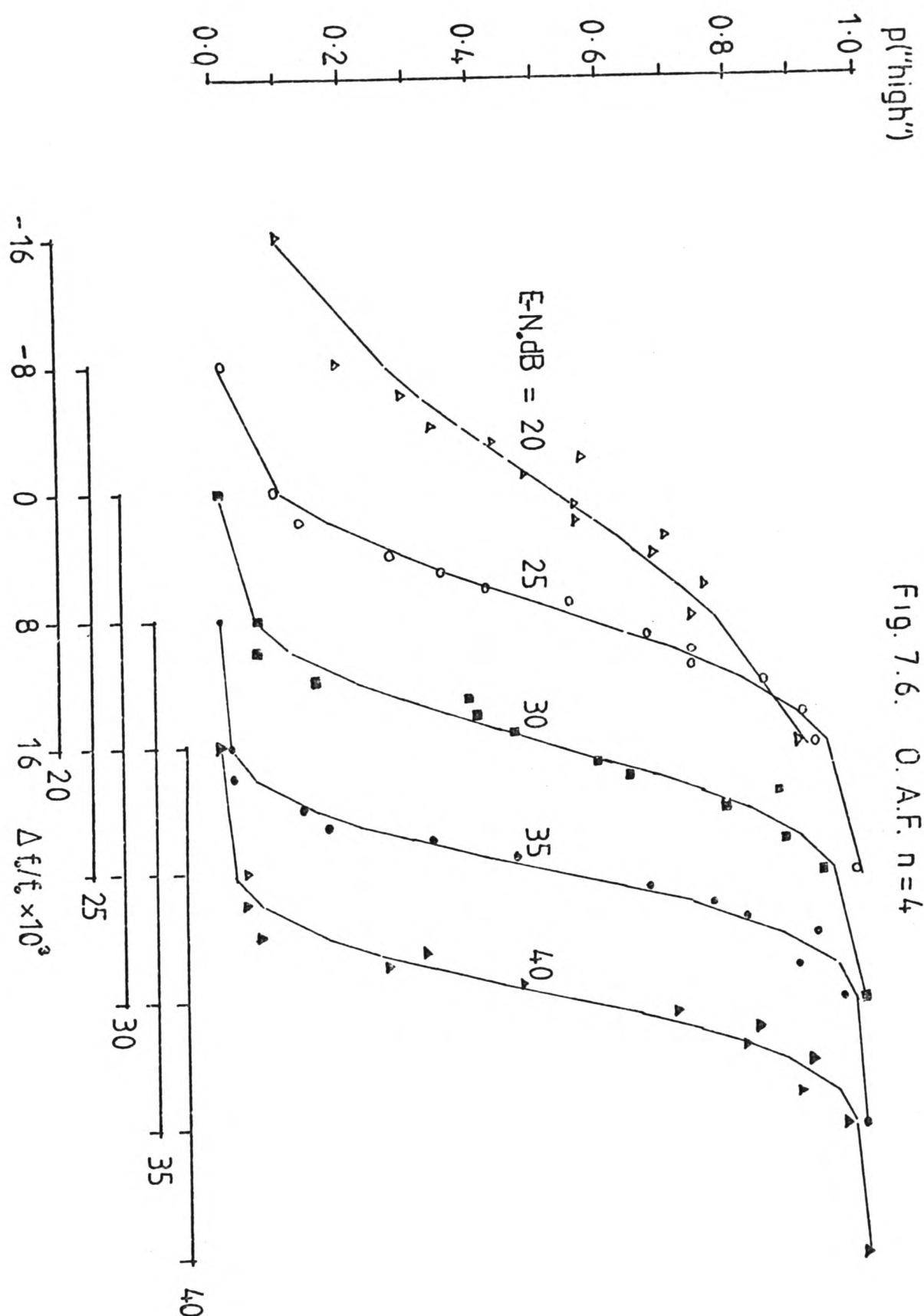
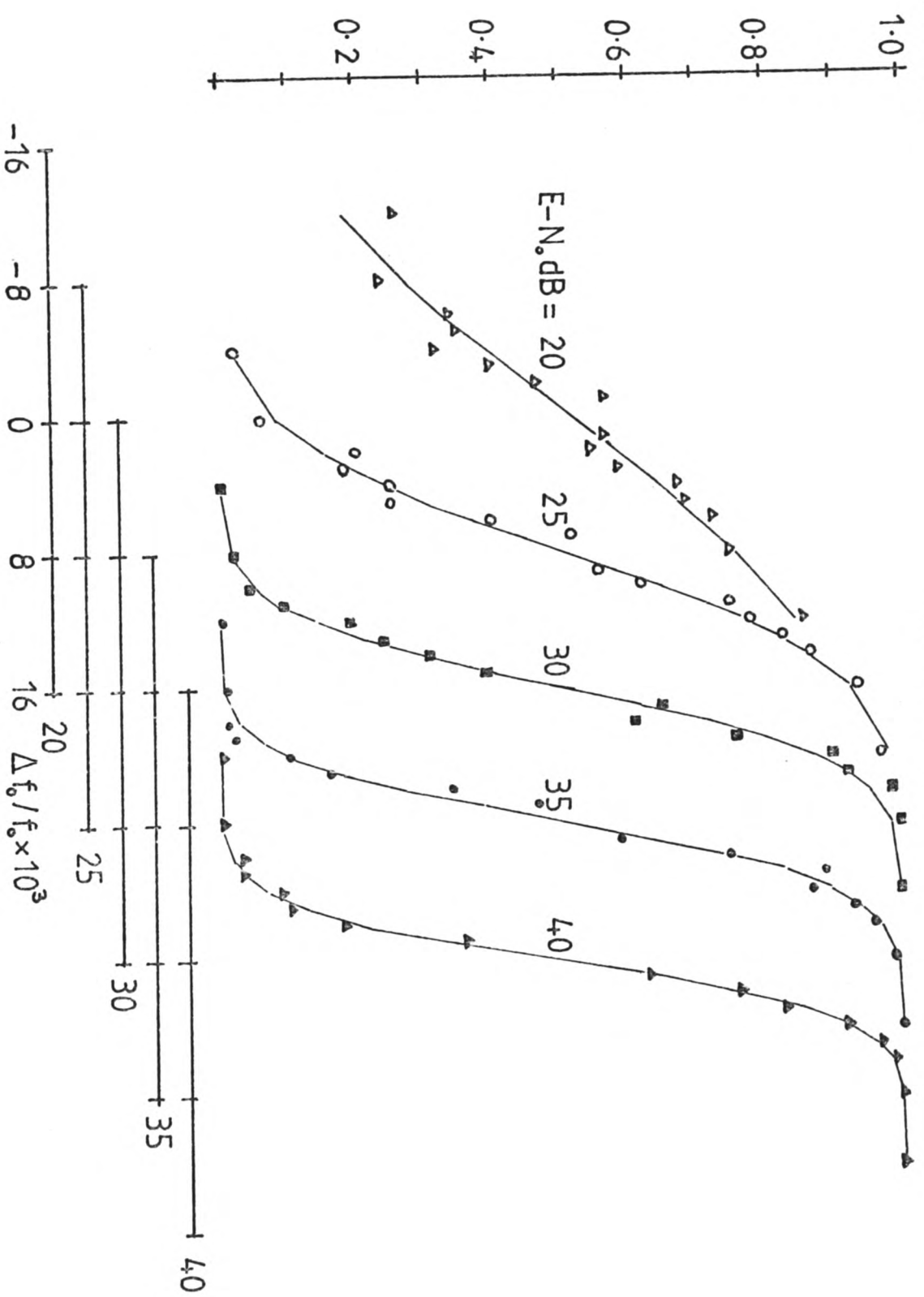


Figure 7.7. Psychometric functions from Experiment 10; pitch discrimination for a two-component complex, $\underline{n} = 4$.
Panel a), observer A.F., 100 trials per point.
Panel b), observer G.S., 100 trials per point except at $E-N_0 = 25$ dB, where there were 150 trials per point. Symbols as figure 7.4.

$p(\text{"high"})$

Fig. 7.7a) 0. A.F. $n=4$



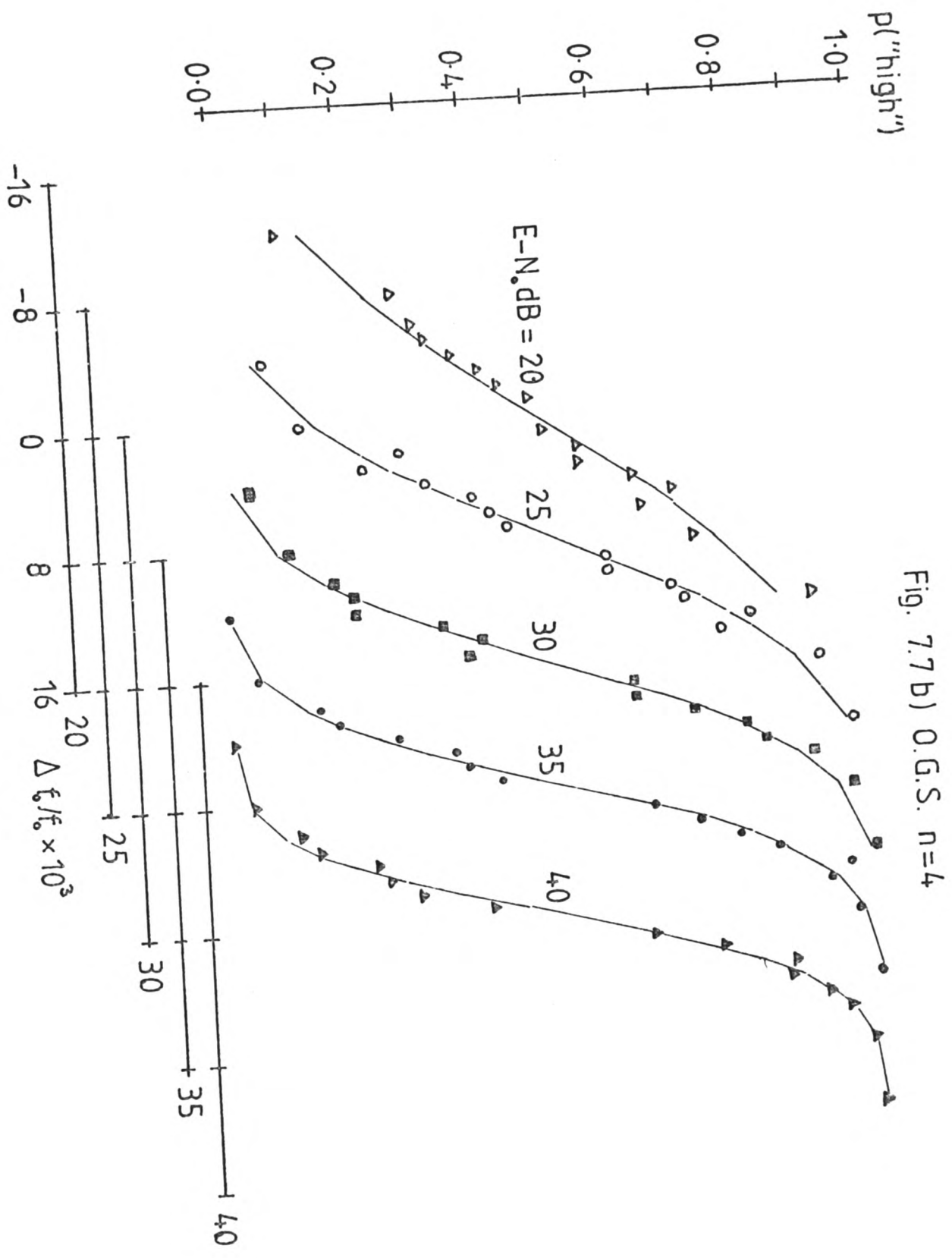


Figure 7.8.Psychometric functions from Experiment 11. Pitch discrimination between the reference signal and a two-component complex, with Δf random within blocks.

Panel a) observer A.F., $\underline{n} = 4$.

Panel b) observer A.F., $\underline{n} = 5$.

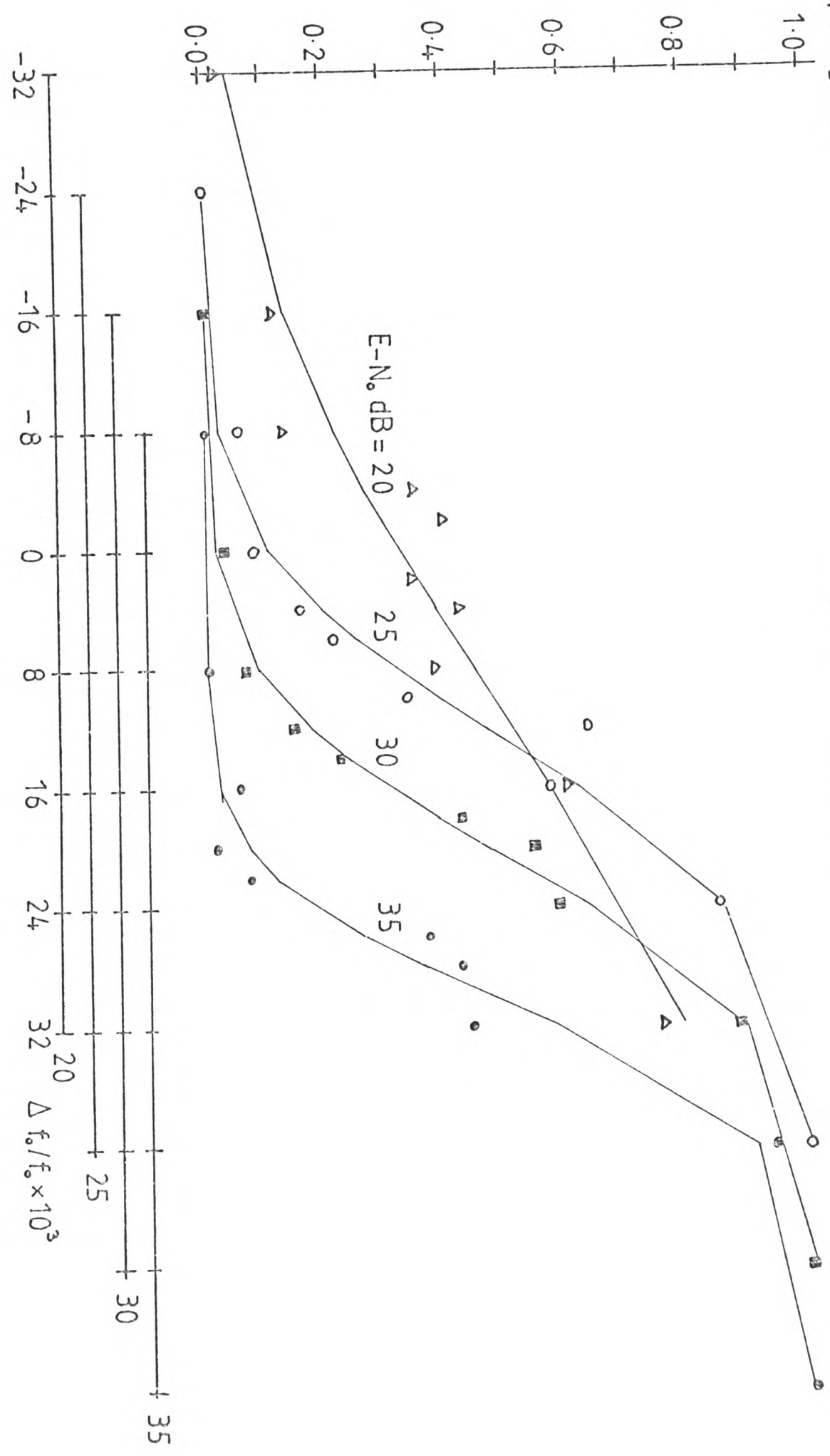
Panel c) observer G.S., $\underline{n} = 4$.

Panel d) observer G.S., $\underline{n} = 5$.

75 trials per point, symbols as figure 7.4.

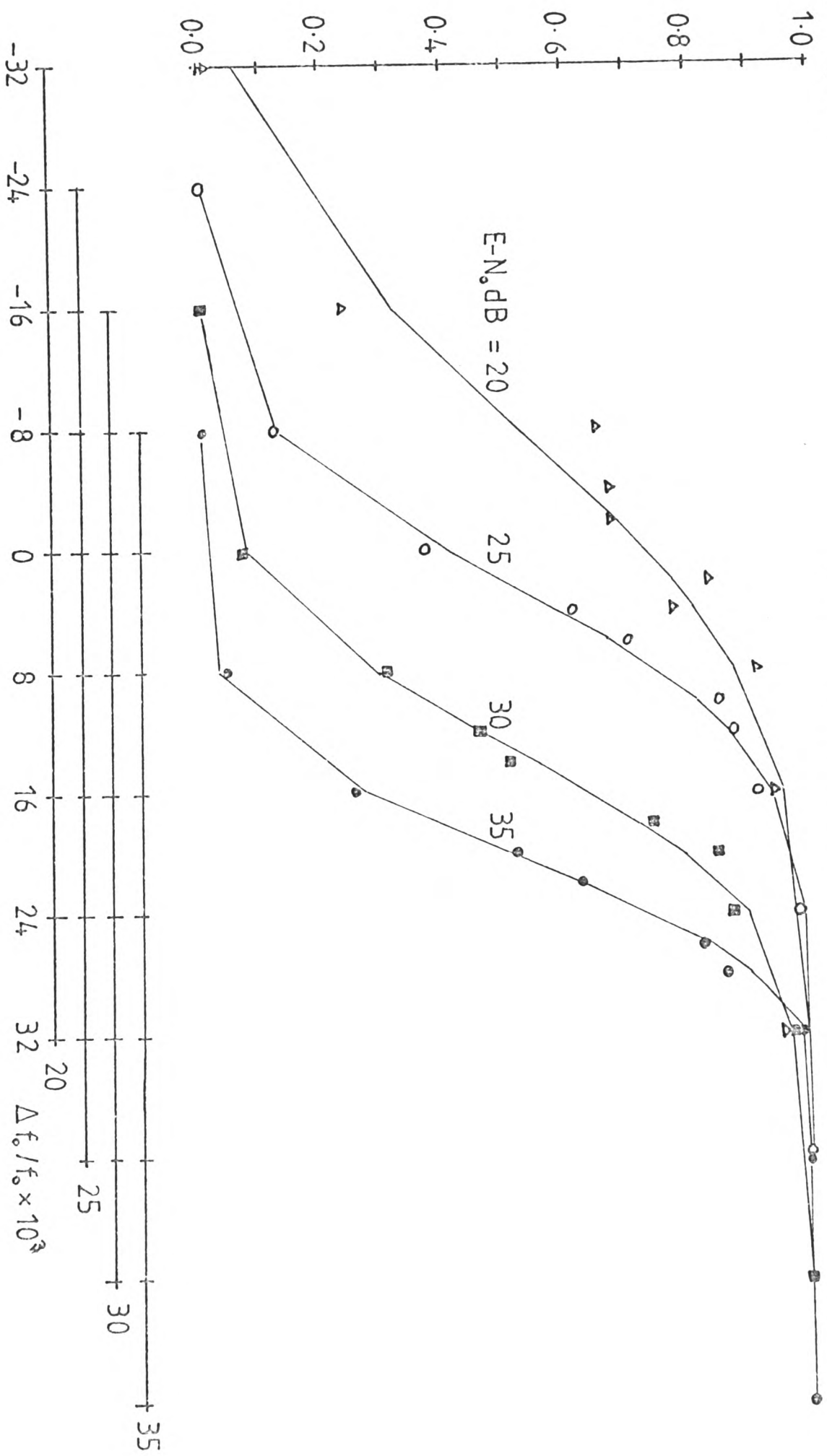
$p(\text{"high"})$

Fig. 7.8a) 0. A.F. $n=4$



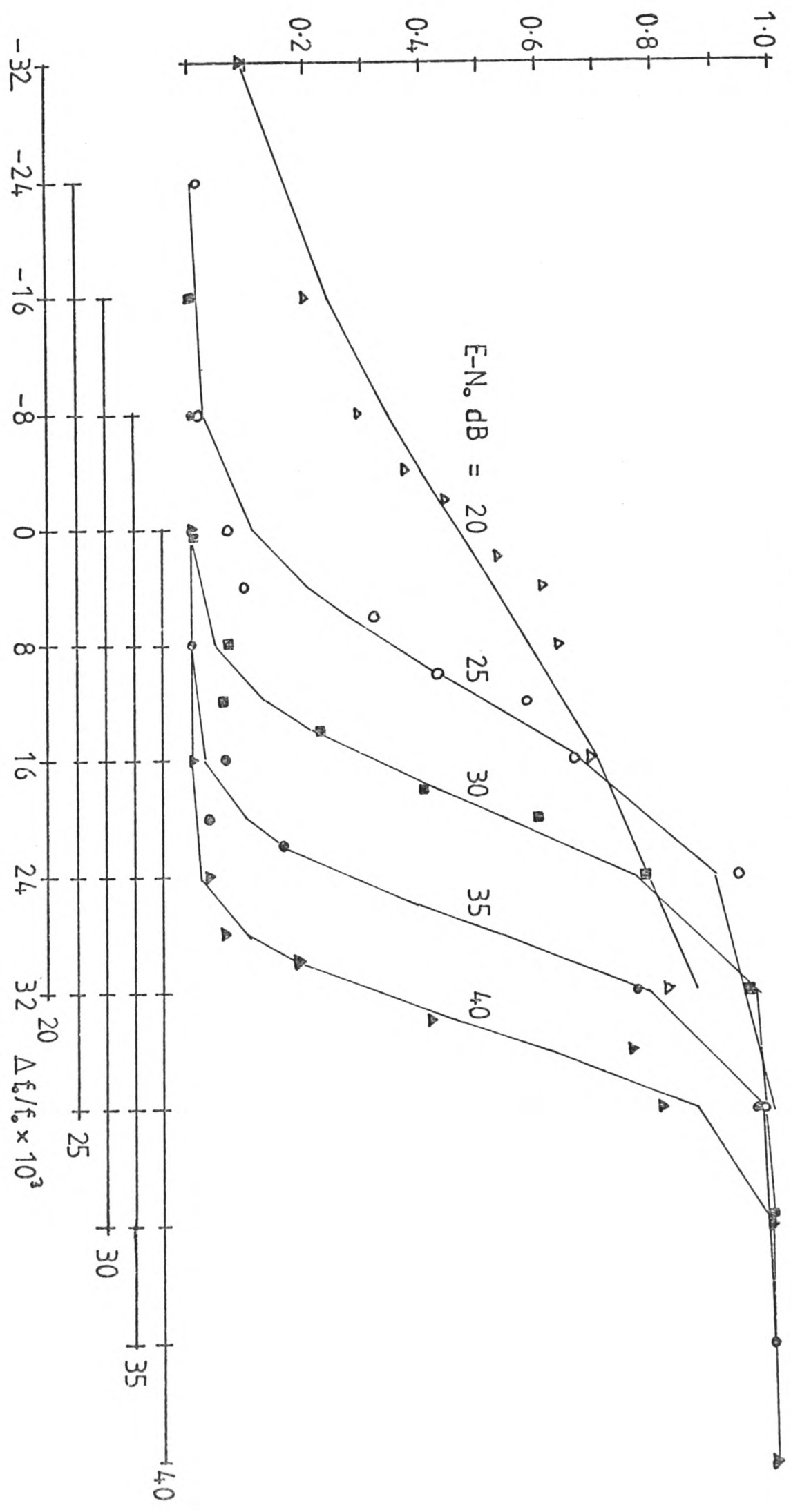
$p(\text{"high"})$

Fig. 7.8b) O.A.F. $n=5$



$p(\text{"high"})$

Fig. 7.8 c) 0.G.S. $n=4$



$P(\text{"high"})$

Fig. 7.8 d) 0. G.S. $n=5$

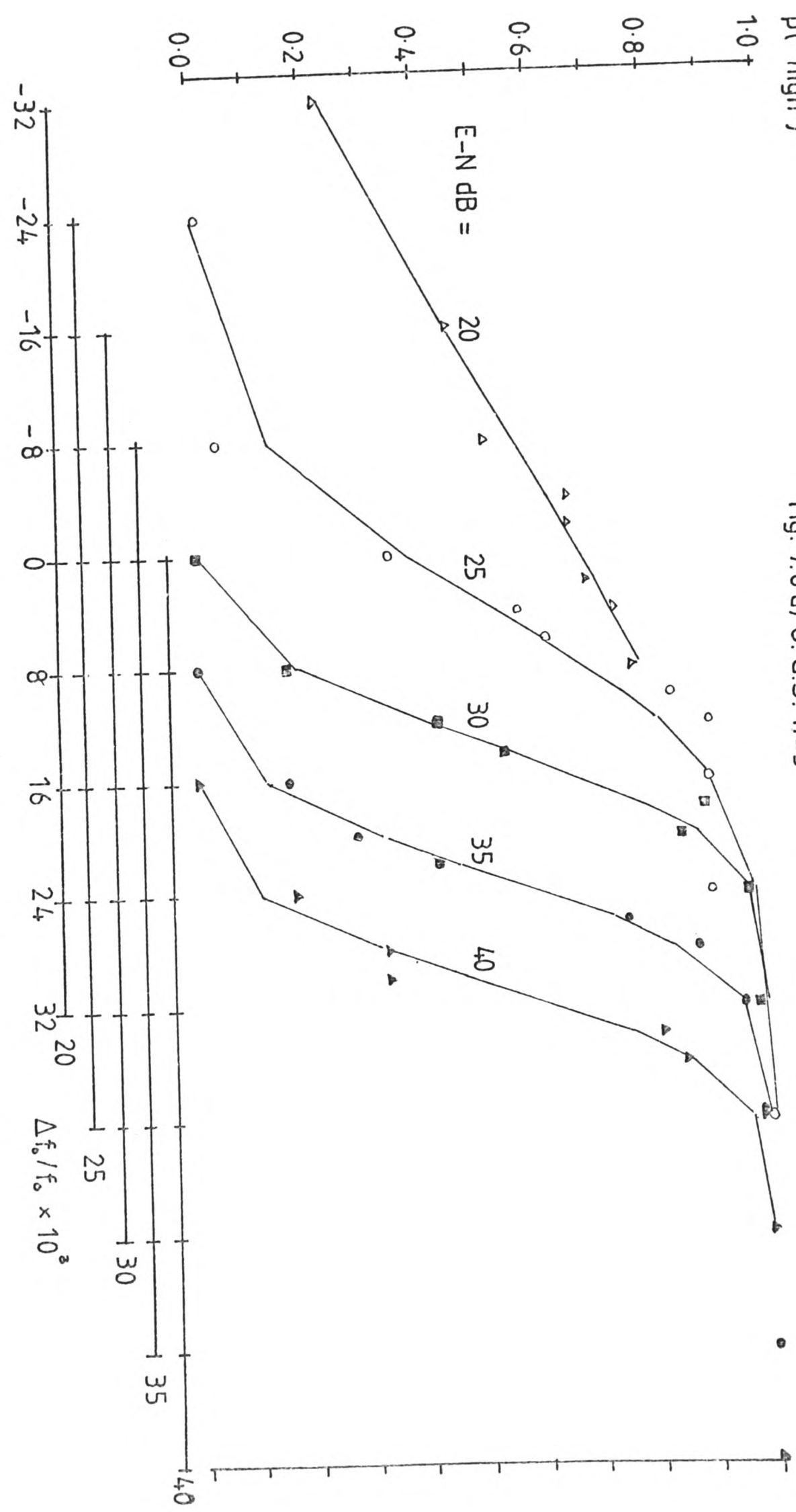


Table 7.1 Probit analysis of observed psychometric functions from Experiment 7 - <u>n</u> is the harmonic number of the test signal. Observer A.F.										
	E-N, dB	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	40	100	168.41	9.04	-0.0015	0.0002	12.617	13	1.0	
	35	100	162.73	8.79	-0.0015	0.0002	9.093	13	1.0	
	30	100	160.85	8.71	-0.0015	0.0003	15.331	13	1.0	
	25	100	109.76	6.42	-0.0016	0.0003	16.272	14	1.0	
	20	100	48.61	3.30	-0.0031	0.0007	21.589*	13	1.0	
n=5	40	100	240.85	12.00	-0.0009	0.0002	9.836	13	1.0	
	35	100	273.22	13.38	-0.0006	0.0002	7.680*	7	1.0	
	30	100	243.19	12.08	-0.0008	0.0002	21.711	13	1.0	
	25	100	167.43	8.99	-0.0012	0.0002	2.393*	10	1.0	
	20	100	82.25	5.04	-0.0006	0.0004	10.737	14	1.0	

Table 7.2. Probit analysis of observed psychometric functions from Experiment 8 - <u>n</u> is the lower harmonic number of the comparison signal. Observer A.F.										
	E-N, dB	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	35	150	161.64	12.31	-0.0018	0.0003	36.368	13	2.798	<0.01
	30	150	155.66	15.76	-0.0018	0.0005	48.053*	10	4.805	<0.01
	25	150	115.36	5.68	-0.0037	0.0003	21.871	14	1.0	
	20	150	78.04	3.87	-0.0057	0.0004	21.598	14	1.0	
n=5	35	150	202.45	8.68	+0.0017	0.0002	18.488	13	1.0	
	30	150	194.11	12.33	+0.0009	0.0003	21.738*	10	2.174	<0.025
	25	150	181.84	8.08	+0.0031	0.0002	11.271*	9	1.0	
	20	150	99.45	4.84	+0.0053	0.0004	7.523*	11	1.0	

Observer G.S.

	E-N, dB	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	40	150	224.62	12.23	-0.0005	0.0002	22.792	13	1.753	<0.05
	35	150	206.67	12.29	-0.0013	0.0002	17.950*	9	1.994	<0.05
	30	150	196.79	17.40	-0.0016	0.0004	38.306*	9	4.256	<0.01
	25	150	122.90	8.73	-0.0028	0.0004	30.348	14	2.168	<0.01
	20	75	94.49	6.91	-0.0027	0.0005	12.519*	11	1.0	
n=5	40	150	229.76	16.67	+0.0017	0.0003	39.444	13	3.034	<0.01
	35	150	219.64	9.31	+0.0023	0.0002	13.624	13	1.0	
	30	150	158.542	7.08	+0.0033	0.0002	13.947*	10	1.0	
	25	150	123.28	6.00	+0.0046	0.0003	10.679	14	1.0	
	20	75	65.52	6.11	+0.0068	0.0008	16.389	12	1.0	

Table 7.3. Probit analysis of observed psychometric functions from Experiment 9 - <u>n</u> is the lower harmonic number of the comparison signal. Observer A.F.										
	E-N ₀ dB	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	40	100	266.54	22.02	-0.0004	0.0003	36.981	13	2.845	<0.01
	35	100	275.44	13.44	-0.0002	0.0002	21.946	13	1.0	
	30	100	205.85	10.55	-0.0003	0.0002	22.466	13	1.0	
	25	100	179.97	9.52	-0.0010	0.0002	16.081	13	1.0	
	20	100	85.427	5.21	-0.0007	0.0004	18.099	14	1.0	
	20	100	82.25	5.04	-0.0006	0.0004	10.737	14	1.0	

Table 7.4. Probit analysis of observed psychometric functions from Experiment 10 - <u>n</u> is the harmonic number of the test signal.										
Observer A.F.										
	E-N ₀ dB	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	40	100	358.70	15.75	+0.0001	0.0001	6.132	14	1.0	
	35	100	310.79	13.31	-0.0001	0.0001	15.658	14	1.0	
	30	100	271.12	11.64	-0.0002	0.0002	15.621	14	1.0	
	25	100	174.95	7.92	-0.0003	0.0002	10.021	14	1.0	
	20	100	79.66	5.62	-0.0008	0.0004	11.123	14	1.0	
Observer G.S.										
	E-N ₀ dB	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	40	100	254.68	10.41	-0.0005	0.0002	7.107	14	1.0	
	35	100	224.501	9.73	-0.0003	0.0002	8.550	14	1.0	
	30	100	184.22	8.24	-0.0005	0.0002	13.254	14	1.0	
	25	100	139.81	5.61	-0.0006	0.0002	14.733	14	1.0	
	20	100	90.33	5.78	-0.0001	0.0004	9.234	14	1.0	

Table 7.5. Probit analysis of observed psychometric functions from Experiment 11 - <u>n</u> is the lower harmonic number of the comparison signal. Observer A.F.										
	E-N ₀ dB	N	Slope	S.E.	50% Point	S.E.	χ^2	df	hf	p
n=4	35	75	140.39	18.58	+0.0066	0.0009	23.491	8	2.936	<0.01
	30	75	106.69	8.17	+0.0046	0.0006	5.165	8	1.0	
	25	75	95.84	10.51	+0.0049	0.0009	16.513	8	2.064	<0.05
	20	75	39.39	3.54	+0.0114	0.0015	12.371	8	1.0	
n=5	35	75	156.58	12.02	-0.0042	0.0005	3.567	7	1.0	
	30	75	111.70	8.54	-0.0033	0.0006	3.748	8	1.0	
	25	75	112.58	8.65	-0.0061	0.0006	3.428	8	1.0	
	20	75	67.26	5.05	-0.0094	0.0009	11.063*	5	1.0	

Table 7.5 continued. Observer G.S.										
	E-N ₀ dB	N	Slope	S.E.	50% Point	S.E.	χ ²	df	hf	p
n=4	40	75	203.03	15.99	+0.0025	0.0004	11.178	7	1.0	
	35	75	177.23	13.79	+0.0038	0.0004	14.142	8	1.0	
	30	75	153.25	11.73	+0.0033	0.0004	7.632	8	1.0	
	25	75	105.40	8.07	+0.0038	0.0006	9.552*	5	1.0	
	20	75	38.20	3.41	+0.0024	0.0013	7.02	8	1.0	
n=5	40	75	191.10	14.88	-0.0017	0.0004	11.244	7	1.0	<0.01
	35	75	174.73	13.50	-0.0015	0.0004	2.632*	3	1.0	
	30	75	174.39	13.51	-0.0028	0.0004	5.380*	3	1.0	
	25	75	98.65	12.21	-0.0050	0.0011	15.362*	5	3.072	
	20	75	37.69	4.57	-0.0126	0.0016	1.449	6	1.0	

In the case of Experiments 8, 10 and 11, the two signals on a given trial differed in structure, so that the resulting internal representations cannot be assumed to be of equal variance. The relative standard deviation of each observer's internal representation of the first of these two signals, the reference signal, was taken to be equal to the DL for this signal as estimated in Experiment 4, and the relative standard deviation $\sigma(f'_0)/f_0$ of the residue pitch of the comparison signal was estimated as $\sigma(x'_2) = \sqrt{(\sigma^2(x'_2 - x'_1) - \sigma^2(x'_1))}$, where x'_1 and x'_2 correspond respectively to the hypothetical internal representations of the pitches of the reference and comparison signals, and $\sigma(x'_2 - x'_1)$ is the reciprocal of the slope of the theoretical function. In the case of Experiments 7 and 10, both signals on a given trial differed only by Δf , so that here the assumption of equal variance is reasonable, and the relative standard deviation $\sigma(f')/f$ was estimated as $\sigma(f') = \sigma(x'_2 - x'_1)/\sqrt{2}$.

Estimates of $\sigma(f'_j)/f_j$.

The estimates of $\sigma(f'_j)/f_j$, where $f_j = n(209.4)$ Hz, derived from the performance of observer A.F. in Experiment 8 are shown in figure 7.9, together with the predicted value of $\sigma(f'_0)/f_0$, where this value is taken as

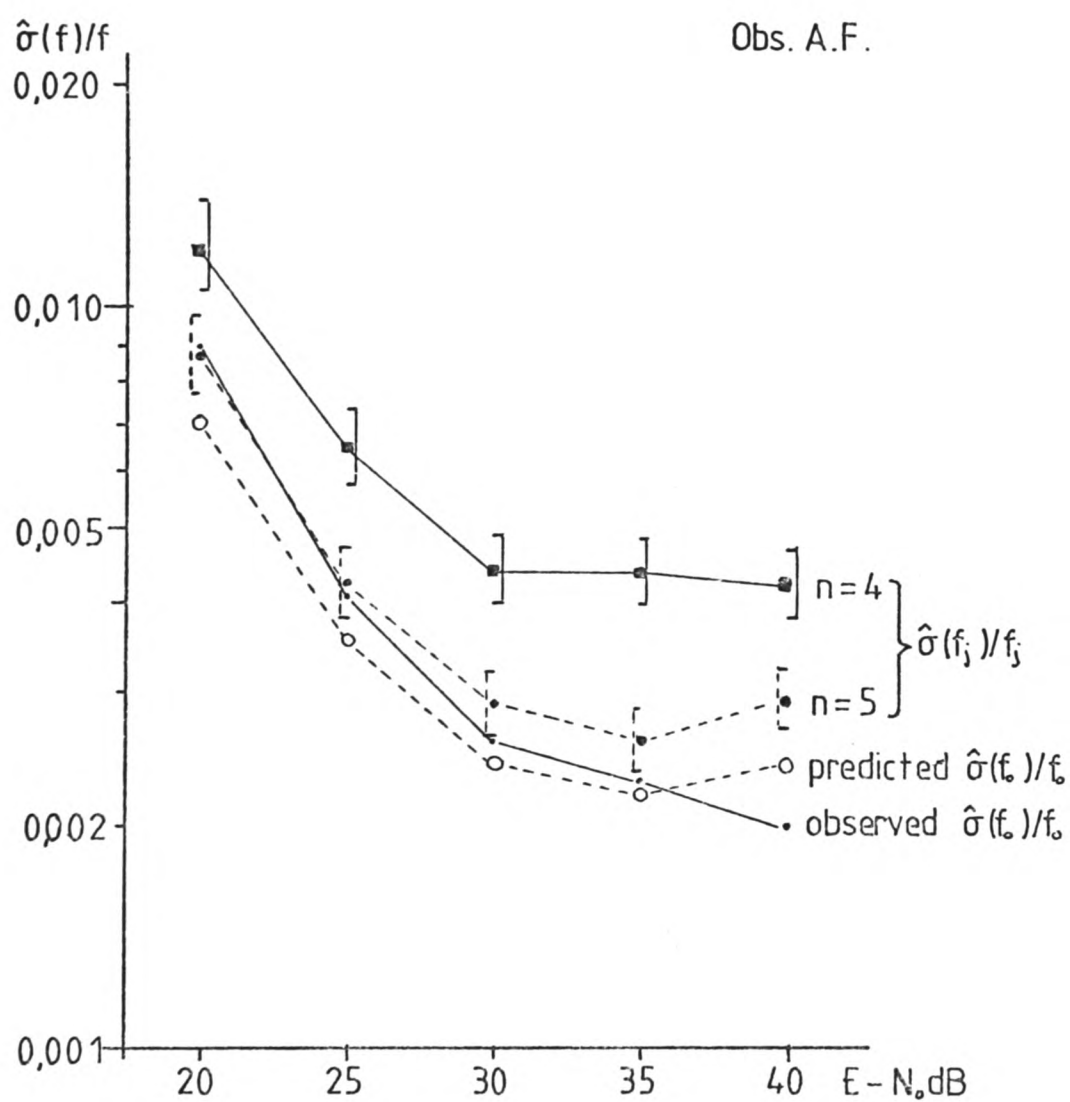
$$\sigma(f'_0)/f_0 = \sqrt{\sum_{j=1}^2 1/[\sigma(f'_j)/f_j]^2}.$$

The estimates of $\sigma(f'_j)/f_j$ for $n=4$ are consistently larger than those for $n=5$, presumably because the low-pass noise spread to the frequency selective channel responding to this component, resulting in a smaller effective signal-to-noise ratio.

Estimates of $\sigma(f'_0)/f_0$ and the Theoretical Predictions.

The estimated values of $\sigma(f'_0)/f_0$ as a function of $E-N_0$ for Experiments 8 and 10 for each observer are shown in figure 7.10. The results from observer A.F. in Experiment 10 are also shown in figure 7.9 together with the predicted values of $\sigma(f'_0)/f_0$ derived from the estimates of $\sigma(f'_j)/f_j$ from Experiment 7. The estimates of $\sigma(f'_0)/f_0$ for this observer from Experiment 10, where stimulus uncertainty is minimal, are very close to these predictions. This result thus replicates the correspondence found between the predicted and estimated values of $\sigma(f'_0)/f_0$ from Experiments 3 and 6, which were similar to Experiments 7 and 10 apart from the independent variable, which in the earlier experiments was the presentation

Figure 7.9. Estimates of $\sigma(f'_j)/f_j$ as a function of $E-N_0$, based on the performance of observer A.F. in Experiment 7. Also shown are the estimates from the same observer and Experiment 10 of $\sigma(f'_0)/f_0$ and the predicted value of $\sigma(f'_0)/f_0$ from the spectral-pattern models based on the estimates of $\sigma(f'_j)/f_j$. The prediction of the periodicity theory is that $\sigma(f'_0)/f_0$ should not be smaller than the minimum of $\sigma(f'_j)/f_j$.



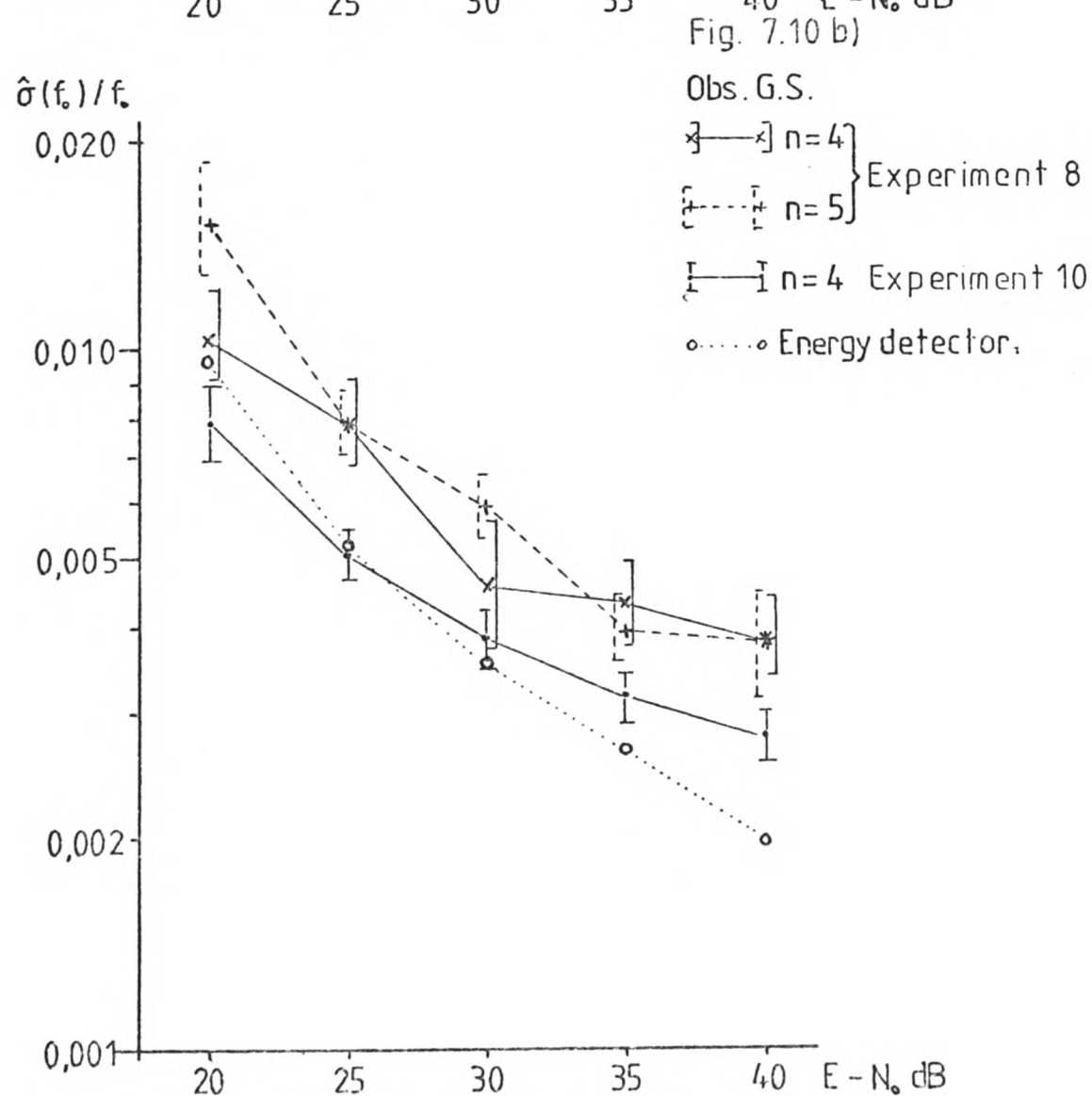
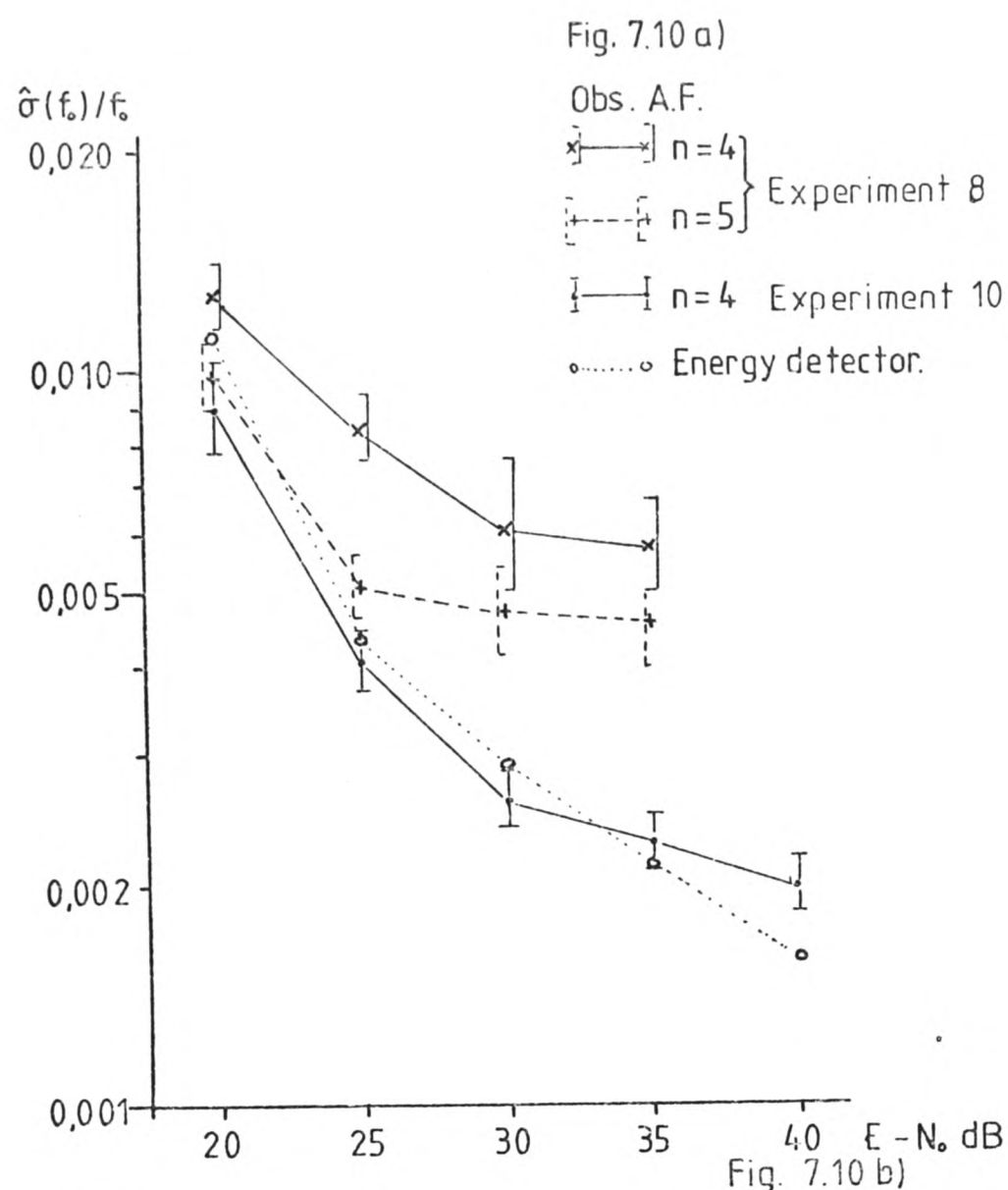


Figure 7.10. Estimates of $\sigma(f_0)/f_0$ from Experiments 8 and 10. Panel a), observer A.F., panel b), observer G.S. The error bars represent 95% confidence intervals. Also shown are the predictions for the estimates from Experiment 10 from the Energy Detector model: the model parameters giving these predictions are given in the text (table 7.7).

level of the components. Because the estimates of $\sigma(f'_j)/f_j$ from Experiment 7 are not equal for both components, the predicted value of $\sigma(f'_0)/f_0$ is also very similar to the smaller estimate of $\sigma(f'_j)/f_j$, which here is for $n=5$, and it is thus possible that the performance of observer A.F. in Experiment 10 was based upon information from this component alone, rather than upon a joint perception of the two components. However, this explanation was not viable for the results of Experiment 6, where the estimated $\sigma(f'_0)/f_0$ was consistently smaller than the smaller estimate of $\sigma(f'_j)/f_j$ from Experiment 3. It is, therefore, reasonable to accept that the results of Experiment 10 do represent performance based upon a joint perception of the two components. Because the variation between the estimates of $\sigma(f'_j)/f_j$ at the signal-to-noise ratios used here is greater than that achieved in Experiment 3, the correspondence of the predicted and estimated values of $\sigma(f'_0)/f_0$ obtained here represents a more general demonstration of the validity of the relationship between $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ predicted by the spectral-pattern recognition models.

The Effect of Frequency Uncertainty.

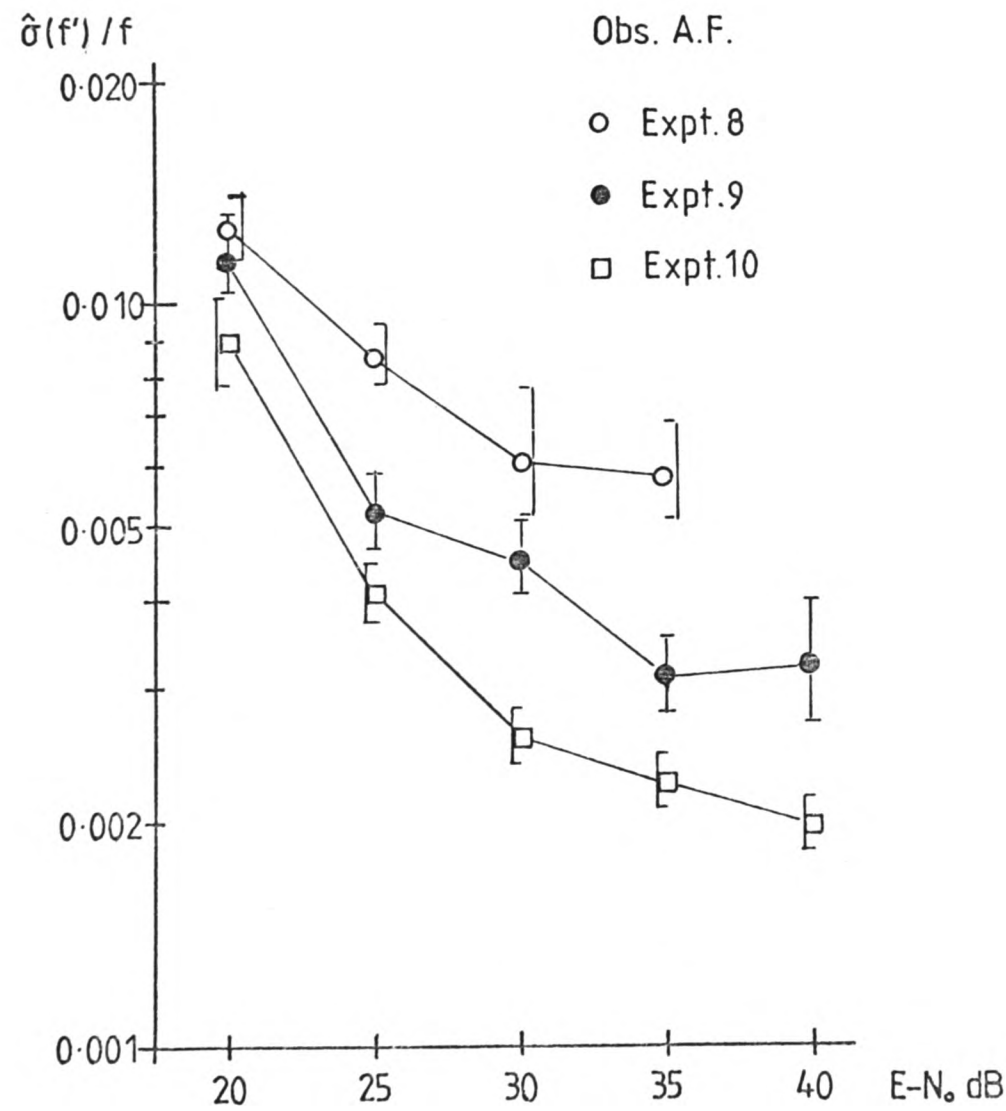
The difference between the estimates from Experiments 8 and 9 should represent the effect of frequency uncertainty upon $\sigma(f'_0)/f_0$, since in Experiment 8, the comparison signal

contained two of three possible components, while in Experiment 9, only two components could occur. The estimates from observer A.F. of $\sigma(f'_0)/f_0$ from Experiment 8 for the comparison signal containing the 4th and 5th harmonics are larger than those from Experiment 9 by an average factor of 1.490. These estimates, and those from the same observer in Experiment 10, are displayed together in figure 7.11. The uncertainty difference would lead to a predicted factor of 1.22, and it seems that the observed factor does not differ to a significant degree from the theoretical value. It is thus concluded that stimulus uncertainty does affect pitch discrimination performance, and that the degree of that effect is approximately consistent with the predictions of a multiple-band model of frequency coding.

Evidence for an Internal Contribution to $\sigma(f'_0)/f_0$.

In the previous chapter, it was proposed that such an effect of uncertainty was responsible for the difference in pitch discrimination performance between two tasks which differed with respect to stimulus uncertainty. There was also a second difference between these tasks; the task with no frequency uncertainty (Experiment 6) involved the comparison of the pitches of two signals composed of the same harmonics of slightly different fundamental frequencies, while the task involving uncertainty (Experiment 5) required the comparison of the pitches of two signals with no common harmonics. Those two

Figure 7.11. A comparison of the estimates of $\frac{\sigma(f'_0)}{f_0}$ from observer A.F. in Experiments 8, 9 and 10. The error bars represent 95% confidence intervals.



tasks have also been presented here as Experiments 8 and 10. The estimates of $\sigma(f'_0)/f_0$ from these two experiments - considering for Experiment 8 only the comparison signal containing the 4th and 5th harmonics - differ by an average factor of 2.126 for observer A.F. and 1.380 for observer G.S., the estimates from Experiment 8 being the larger. A factor of 2.126 is rather too large to be due entirely to the uncertainty difference, so that it seems likely that some additional internal noise process contributed to the performance of observer A.F. This conclusion is supported by the differences in the estimates of $\sigma(f'_0)/f_0$ for this observer from Experiments 9 and 10. Here there was no uncertainty difference, while the comparison in Experiment 10, but not in Experiment 9, could be made within the same frequency-selective channels. The estimates from this observer in Experiment 9 exceed those from Experiment 10 by an average factor of 1.490, and from the confidence limits shown on figure 7.11, this is a significant difference.

The factor of 1.38 shown by the estimates from Experiments 8 and 10 from observer G.S. is however similar to the factor of 1.22 expected from the uncertainty difference, so that for this observer, uncertainty is probably sufficient to explain this difference. Alternatively, it might be the case that this observer's performance is unaffected by frequency uncertainty, and that the differences between his performance in the two

tasks considered here is entirely due to an internal noise which affects only those pitch comparisons made between frequency-selective channels. The assumption of such an internal noise is discussed further in the next chapter, where general conclusions are drawn from all of the pitch discrimination experiments described here.

The Importance of Between-Trial Comparisons.

There are only rather small differences between the estimates of $\sigma(f'_0)/f_0$ derived from Experiments 8 and 11, which differed in that the value of $|\Delta f|$ was constant within each block for Experiment 8, while in Experiment 11, five different values of $|\Delta f|$ were presented in each block of trials. The estimates of $\sigma(f'_0)/f_0$ from these two experiments appear together in figure 7.12. If, as was suggested earlier, between-trial comparisons of the residue pitches or component frequencies of the comparison signals were primarily responsible for the observed performance, this procedural difference should reduce performance considerably. The increase in the estimated value of $\sigma(f'_0)/f_0$ between Experiments 8 and 11 was equivalent to an average factor of 1.37 for observer G.S., and a factor of 1.53 for observer A.F.; this order of performance decrement is not consistent with the hypothesis that between trial comparisons play any major role. A decrement of this order is however comparable with the performance decrement which might be

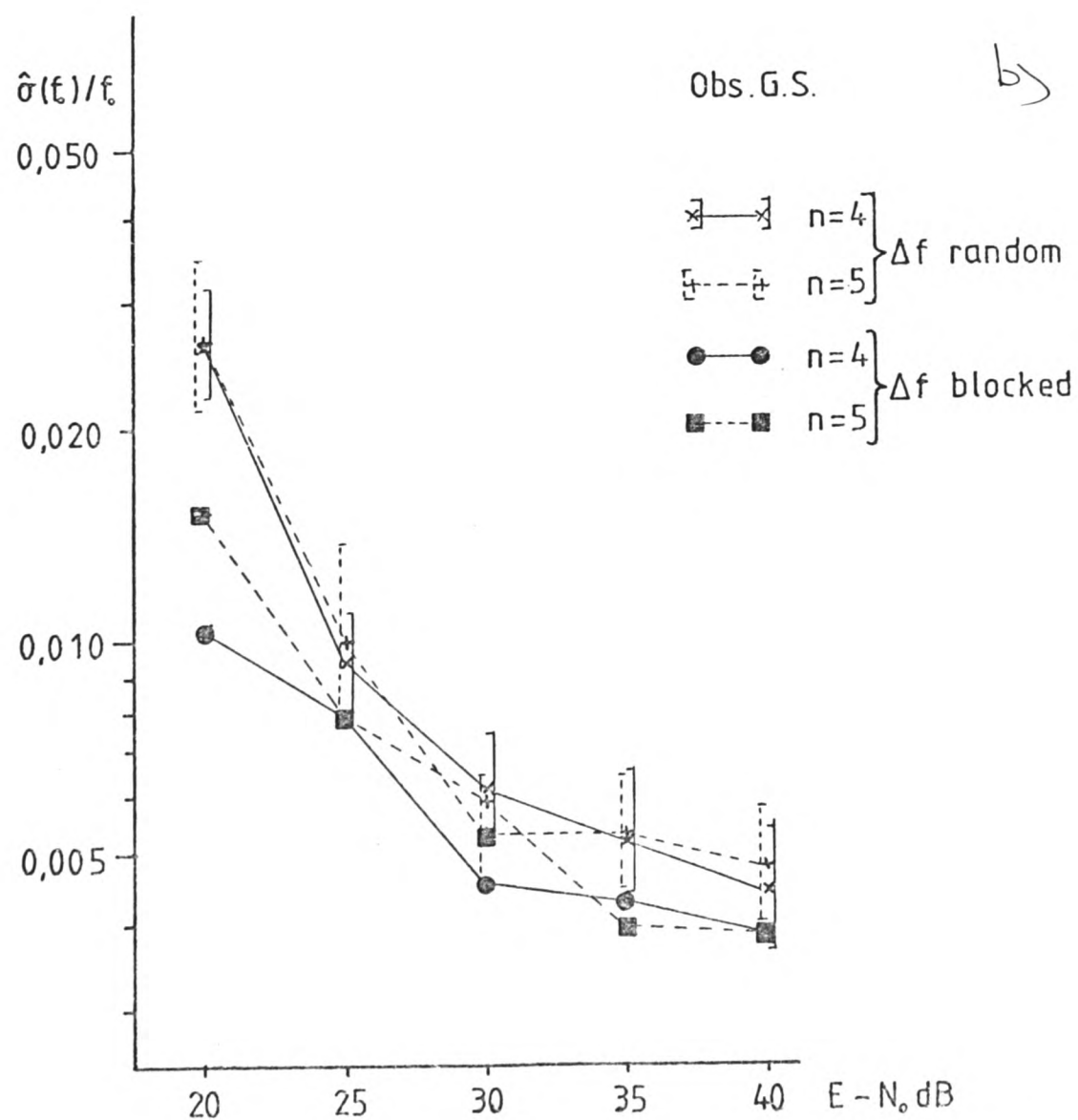
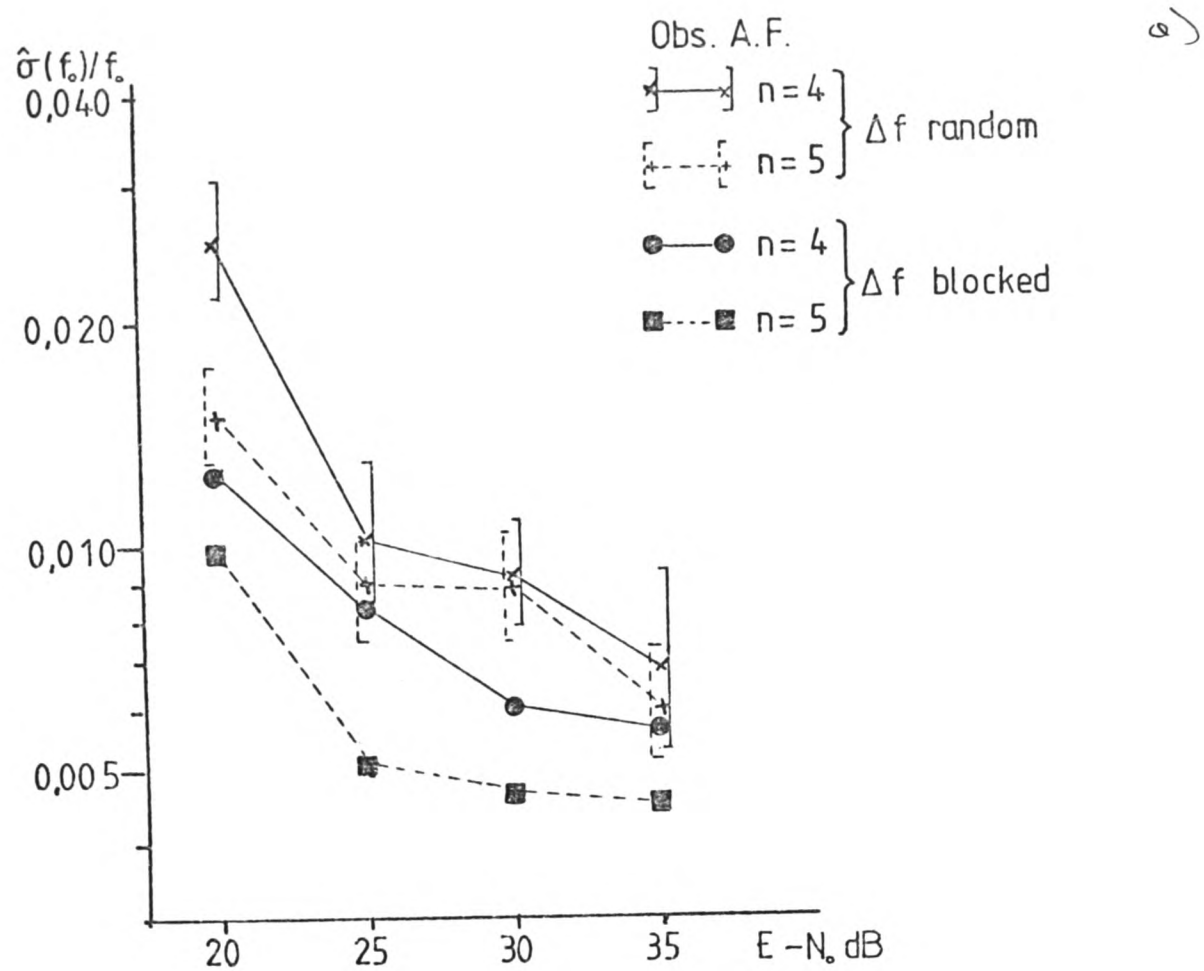


Figure 7.12. A comparison of the estimates of $\hat{\sigma}(f_o)/f_o$ from Experiments 8 and 11: the effect of randomizing Δf within blocks. Panel a) observer A.F., panel b), observer G.S.

expected if the larger number of pitch differences led to an increase in the variability of the pitch difference criterion proposed in the decision model described earlier. There is, however, one difficulty with this explanation. A larger criterion variance in this task should lead to a constant additive difference between the squares of the estimates of $\sigma(f'_0)/f_0$ from Experiments 8 and 11: The observed differences, which are illustrated in figure 7.13, increase with $\sigma(f'_0)/f_0$. This may indicate that the effect of randomizing Δf is to increase $\sigma^2(f'_0)$ by a multiplicative constant. An alternative possibility is that there is indeed an added internal variance, whose effect is obscured by the greater variability of the larger estimates of $\sigma(f'_0)$.

The Dependence of Discrimination on Signal-to-Noise Ratio.

The estimates of $\sigma(f'_0)/f_0$ as a function of $E-N_0$ derived from Experiment 10 can be reasonably well described by the Energy Detector model. The observed DLs and the predictions of this model are displayed in figure 7.10. The model parameters were derived in the manner described in Chapter 3, except that the model was set up with filters around both 838 and 1047 Hz, and these parameters are presented in table 7.6.

Figure 7.13. The square root of the difference between the squared estimates of $\sigma(f'_0)/f_0$ from Experiments 11 and 8. The point marked with an asterisk is the square root of the absolute value of a negative difference.

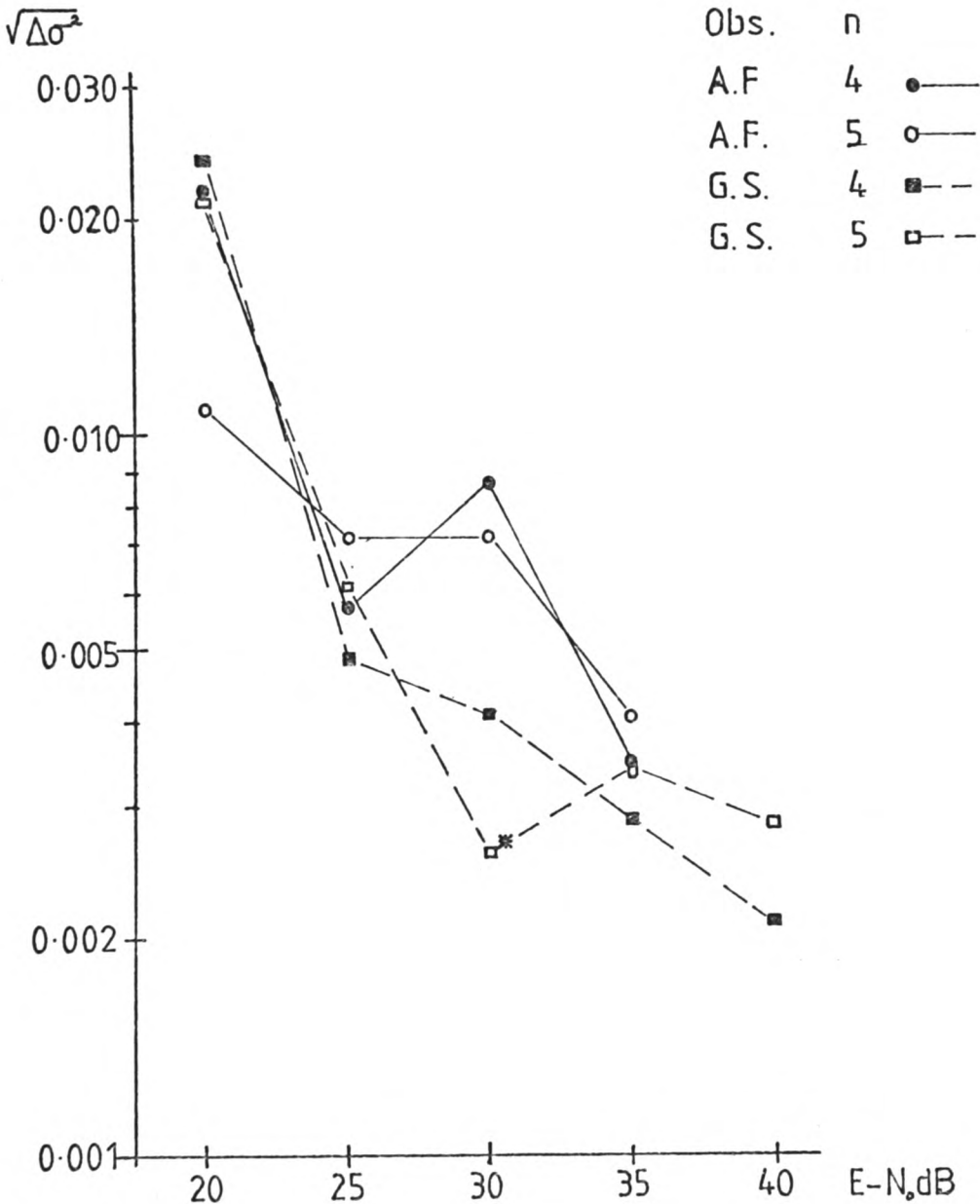


Table 7.6. Parameters of the Energy Detector model found to fit the estimates of $\sigma(f'_0)/f_0$ as a function of $E-N_0$ from experiment 10.

Observer	Filter width (Hz)	Integration time (sec).
A.F.	45.227	0.2794
G.S.	61.091	0.1038

The quoted filter width is the standard deviation of a Gaussian filter centred at around 1047 Hz. The filter width at about 838 Hz was 4/5 of the quoted value. The equivalent rectangular bandwidth of such a filter is $\sqrt{2\pi}$ larger than the quoted width.

The filter width parameters fitted to these data are consistent with conventional estimates of the critical bandwidth, the equivalent rectangular bandwidths of the fitted filters at 1000 Hz being between about 110 and 150 Hz. The DLs predicted by the model are tolerably close to those observed, which show a slope of about $1/\sqrt{N_0}$ for observer A.F., and a somewhat smaller slope for observer G.S. It seems then, that the Energy Detector model is sufficient to describe the dependence of pitch discrimination for these complex signals upon signal-to-noise ratio. With due allowance for effects of frequency uncertainty and an additional internal noise where pitch comparisons are made between frequency-selective channels, the form of the signal-to-noise ratio dependence of the data from the other tasks described here is also consistent with the predictions of this model.

Summary.

A number of experiments investigating the dependence of

pitch discrimination for two-component harmonic signals and frequency discrimination for the components of those signals upon signal-to-noise ratio have been described, and the following main conclusions have been drawn.

A/ The relative pitch difference limen for a two-component complex signal, measured in a 2AFC task where the two signals differed only in fundamental frequency, can be well described by statistical summation of the relative frequency difference limens for the two components of that signal, where the latter relative DLs were obtained from a 2AFC task where the second component was present as a continuous tone. These latter relative difference limens may be equated with the parameter $\sigma(f'_j)/f_j$ in the spectral-pattern recognition models of pitch, while the relative residue pitch DL for the two-component complex may be equated with the parameter $\sigma(f'_0)/f_0$. Those models, and also a multiple-band model of frequency discrimination, lead to the prediction that $1/(\sigma(f'_0)/f_0)^2$ is equal to $\sum_{j=1}^2 1/(\sigma(f'_j)/f_j)^2$, and this prediction is supported by the observed performance.

B/ The pitch difference limen for the residue pitch of a two-component harmonic signal appears to depend upon the degree of stimulus uncertainty. Estimates of the relative standard deviation of the internal representation of the residue pitch of such a signal from a 2AFC pitch discrimination task were larger

by a factor of 1.49 where the two-component signal could be composed of two out of three possible components compared to a similar task where only two components were possible. The factor expected from a multiple-band model of frequency coding would be $\sqrt{3/2}$ or 1.22, and the observed factor does not differ significantly from the theoretical value.

C/ Performance in pitch discrimination tasks involving comparisons between two pitches derived from signals having components in different frequency regions is worse than that found in tasks where two pitches are derived from signals having components in the same frequency regions. Estimates of the parameter $\sigma(f'_0)/f_0$ from tasks of the former kind exceed those from a task of the latter kind by a factor of between about 1.5 and 2, and such differences may be explained by the assumption that pitch comparisons based upon information from different frequency-selective channels are affected by an internal noise process which does not affect comparisons made within the same channels. Such a noise process is equivalent to that assumed by Goldstein (1973) to account for performance in tasks where melodies and musical interval were to be identified on the basis of the residue pitches of two-component harmonic signals.

CHAPTER 8

Abstract.

This chapter contains a description of two experiments intended to estimate the effects of signal duration upon the standard deviation of the internal representation of the residue pitch of a two-component harmonic signal. These experiments, Experiments 12 and 13, were essentially similar to Experiments 5 and 6, which were described in Chapter 5. The signal components were equal in amplitude to a 100 msec sinusoidal tone burst at approximately 30 dB SL, and were presented in a background of noise whose spectral-density was 30 dB below that amplitude. The choice of signal duration as an independent variable permits a test of the predictions of the models of frequency discrimination discussed in Chapter 3 for the case of residue pitch. The manipulation of this parameter also allows a comparison of the effects of signal-to-noise ratio with those of signal duration. Since the Energy Detector model predicts that frequency discrimination is primarily determined by the ratio between the signal and noise energies, and signal energy is proportional to signal duration, comparable signal-to-noise energy ratios should result in comparable performance, and thus signal power (or the inverse of noise power) and signal duration

can be traded for each other.

While it has been possible to vary $\sigma(f'_j)/f_j$ by manipulating the level of a background noise, the range of the resulting values of $\sigma(f'_j)/f_j$ only encompassed a factor of about four, and it was hoped that the manipulation of signal duration might result in a greater range of values of this parameter, so that the planned studies of the relationship between $\sigma(f'_j)/f_j$ and the ambiguity of residue pitch could be carried out with stimulus conditions expected to lead to an more easily detectable effect.

The observed duration dependence of $\sigma(f'_0)/f_0$ between 80 and 25 msec was consistent with an inverse dependence upon the square root of duration, while the increase in $\sigma(f'_0)/f_0$ between durations of 25 and 20 msec exceeded that predicted by that relationship.

The estimates of $\sigma(f'_0)/f_0$ obtained here were compared with those obtained from Experiments 8 and 10 above, where signal amplitude and duration were constant and noise power was varied. At equivalent signal-to-noise energy ratios, the estimates of $\sigma(f'_0)/f_0$ were larger for the shorter duration signals. These differences are discussed in terms of differences between the rise and fall times of the signals, and in terms of a modified Energy Detector model with an additive internal variance.

Introductory Comments.

If the residue pitch of a two-component complex signal is derived from independent internal representations of the frequencies of the two components, the models of frequency discrimination described in Chapter 3 can predict the relationship between signal duration and the standard deviation of the internal representation of that residue pitch. According to the Energy Detector model, if the noise level is constant and the signal duration T_s is varied, this standard deviation should be a function of $1/\sqrt{T_s}$. The data of Turnbull (1944), Liang and Chistovich (1961), Henning (1970), and Moore (1973a) for single sinusoids are generally consistent with this relationship where T_s is less than about 200 msec and greater than a frequency dependent minimum, which at 1000 Hz is approximately 20 msec. Since the detected energy which this model operates upon is proportional to the square root of signal duration, and the DL is primarily determined by the difference between the detected signal and noise energies, the effects of signal duration should be comparable with the effects of signal-to-noise ratio. Thus, at all signal durations less than the period of time over which energy is integrated and greater than about 20 msec, the DL should be constant for a equivalent difference between the signal and noise energies.

The model described by Siebert (1970), which is based upon the optimal use of the timing information in the firing pattern of the auditory nerves, predicts that this standard deviation should be a function of $1/T_s^{3/2}$. The modification of Siebert's

model described by Goldstein and Srulovicz (1977) predicts, like the Energy Detector model, that the DL will be inversely proportional to the square root of signal duration, but since this model derives its predictions about the effects of signal and noise levels independently from predictions of the effect of signal duration, there is no reason to expect from this model a direct trading relationship between signal duration and the signal energy to noise energy ratio.

A previous attempt to estimate the effect of signal duration upon discrimination performance for the residue pitch of two-component harmonic signal was reported by Cudahy (1975, Note 5). As in the studies reported here, the fundamental frequency of the signals was in the region of 200 Hz, while the lower harmonic was varied between the 4th and the 10th. The experimental method was similar to that used in Experiment 1 above, in that the two signals presented on a trial contained different pairs of harmonics of similar fundamental frequencies. The DLs reported by Cudahy are larger than those reported above, the smallest relative DLs being about 0.010, and the largest, for a 15 msec signal, being greater than 0.500. The great range of these DLs strongly suggests that the observer's performance was greatly influenced by differences between the frequencies of the components of the signals in the same manner as was seen in the results of Experiment 1 above, and for this reason, these data are of no relevance to the issues addressed here.

Experimental Method.

The two experiments described here each employed the 2AFC discrimination task described in Chapter 3, and are generally similar to Experiments 5 and 6 above. The observer's task was to judge to pitch of the second signal presented on each trial as either higher or lower than the pitch of the first signal.

Observers.

The two observers, A.F. and G.S., who had participated in Experiments 4 to 11 also participated in these experiments.

Stimuli - Experiment 12.

The first signal, the reference signal, presented on each trial was the signal presented in Experiment 4, except that the fundamental frequency of this signal was always equal to 209.4 Hz. Each component of this signal was presented at a level 40 dB above the observer's threshold at 1047 Hz in a silent background. During the presentation of the reference signal, no background noise was present. The second signal, the comparison signal, contained either the 4th and the 5th ($n=4$) or the 5th and the 6th ($n=5$) harmonics of $209.4+\Delta f$ or $209.4-\Delta f$ Hz; each of the two components was presented at an amplitude equivalent to about 30 dB SL for a 100 msec signal. The half-power duration of the comparison signals was varied from 20 to 80 msec, while the reference signal was always 100 msec in duration; each of these signals was gated on and off with an exponential

envelope having a rise and fall time of 5 msec (which compares with a rise and fall time of 12.5 msec for the previous experiments). The noise background was introduced after the first 100 msec of the 200 msec inter-stimulus interval, and was removed again 100 msec after the offset of the comparison signal. A single trial is illustrated schematically in figure 8.1.

The noise consisted of the sum of a white noise whose spectral-density was -30 dB relative to the amplitudes of the components of the comparison signal, and a low-pass filtered noise intended to mask any distortion products lying below the components of the comparison signal with $n=4$. These two noise signals were both derived from a single white noise generator. The cut-off frequency for low-pass filtering was at 613 Hz, and the filter slope was 72 dB/octave. The spectral-density of the low-pass filtered noise at 613 Hz was 20 dB below the levels of the components of the comparison signal as measured electrically at the headphone input with a wave analyser (HP 3530A) set for a 1 Hz measuring bandwidth.

Procedure.

Trials were presented in blocks of 100 with a constant value of $|\Delta f|$. The value of $|\Delta f|$ ranged from 209.4×0.06498 Hz to 209.4×0.00101 Hz. except for the 20 msec comparison signal, where the smallest value of $|\Delta f|$ was 209.4×0.00203 Hz. Within a block, the probability that the comparison signal had a higher fundamental frequency than the reference signal was 0.50,

Figure 8.1. The time-course of a single trial for Experiment 12.

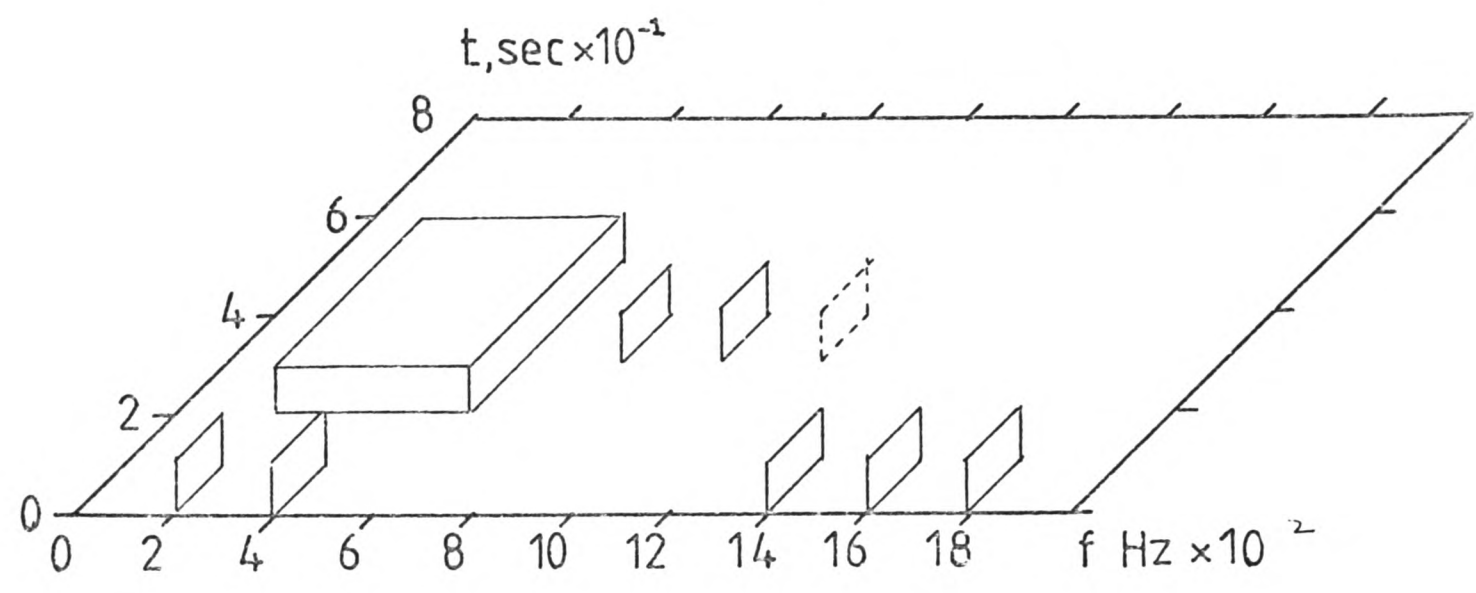
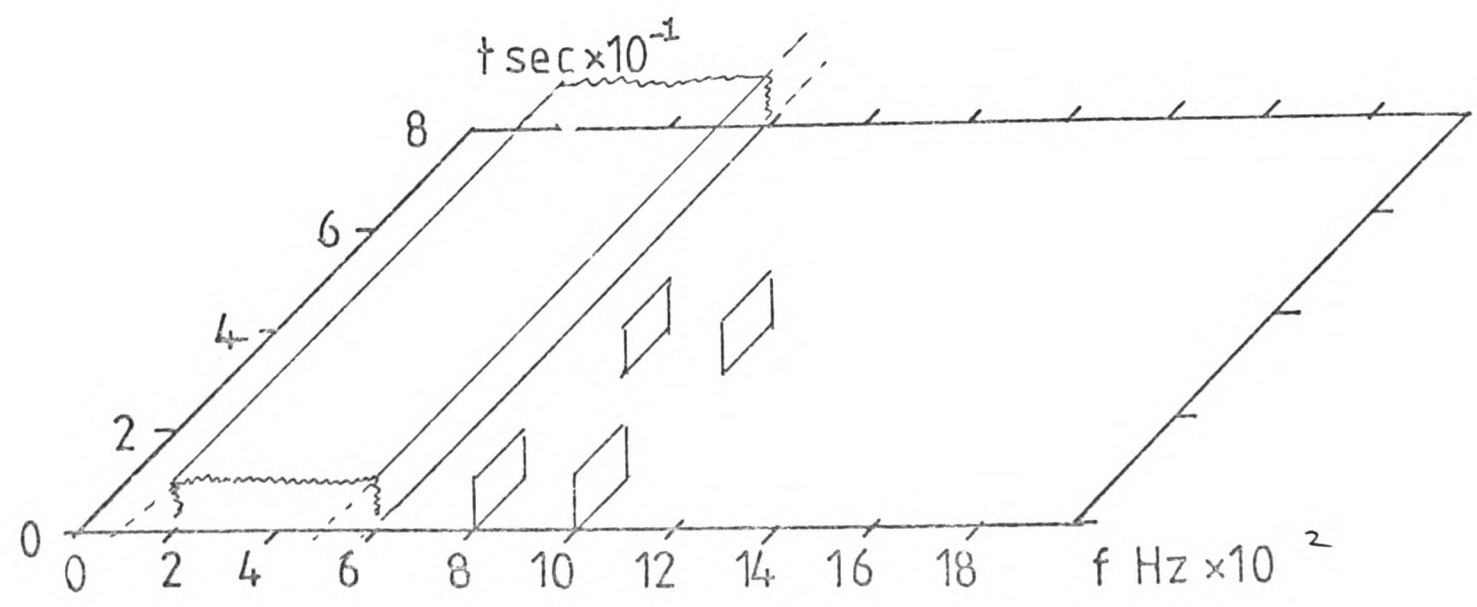


Figure 8.2. The time-course of a single trial for Experiment 13.



the probability that the comparison signal contained the 4th and 5th harmonics was 0.50, and these two events were independent. A series of eight blocks with a constant signal duration was presented in a test session of about one hour in length. Since the previous experiments had shown no effect of the order in which Δf changed over blocks, the value of $|\Delta f|$ always decreased over successive blocks. The order of test sessions at the different signal durations was counterbalanced over the days of presentation. Each observer received at least three series of blocks at each signal duration, and thus received at least 150 trials for each of the two comparison signals at each value of $|\Delta f|$ at each of the signal durations.

Stimuli - Experiment 13.

This experiment was similar to Experiment 12 above, except that both of the stimuli presented on each trial contained the 4th and 5th harmonics of $209.4 + \Delta f/2$ or $209.4 - \Delta f/2$ Hz. The two components of each signal were presented at an amplitude equivalent to about 30 dB SL for a 100 msec duration signal. The signals were presented against a continuous background of white and low-pass filtered noise. The sequence of events on a single trial is illustrated in figure 8.2.

Procedure.

Each block contained 50 trials, and the probability that the second signal had the higher fundamental frequency was 0.50. The value of $|\Delta f|$ ranged from 209.4×0.06496 Hz to $209.4 \times$

0.00101 Hz. Each test session lasted about an hour, and comprised two series of eight blocks, the signal duration being constant within each series. A single stimulus tape was prepared, which was replayed in both directions so that the value of $|\Delta f|$ either increased or decreased over the eight blocks of a series. The ordering over test sessions of blocks at different signal durations and with Δf increasing or decreasing was counterbalanced. Each observer received at least 200 trials at each value of $|\Delta f|$ and signal duration.

Results.

Psychometric Functions.

The observed psychometric functions from these 2 experiments were subjected to a Probit analysis (Finney 1971); the observed and fitted functions appear in figures 8.3 and 8.4, while the parameters of the fitted functions are shown in tables 8.1 and 8.2.

Figure 8.3. Psychometric functions from Experiment 12; pitch discrimination between the reference signal and a two-component comparison signal.

Panel a), observer A.F., $\underline{n} = 4$.

Panel b), observer A.F., $\underline{n} = 5$.

Panel c), observer G.S., $\underline{n} = 4$.

Panel d), observer G.S., $\underline{n} = 5$.

The legend \underline{n} represents the lower harmonic number of the comparison signal.

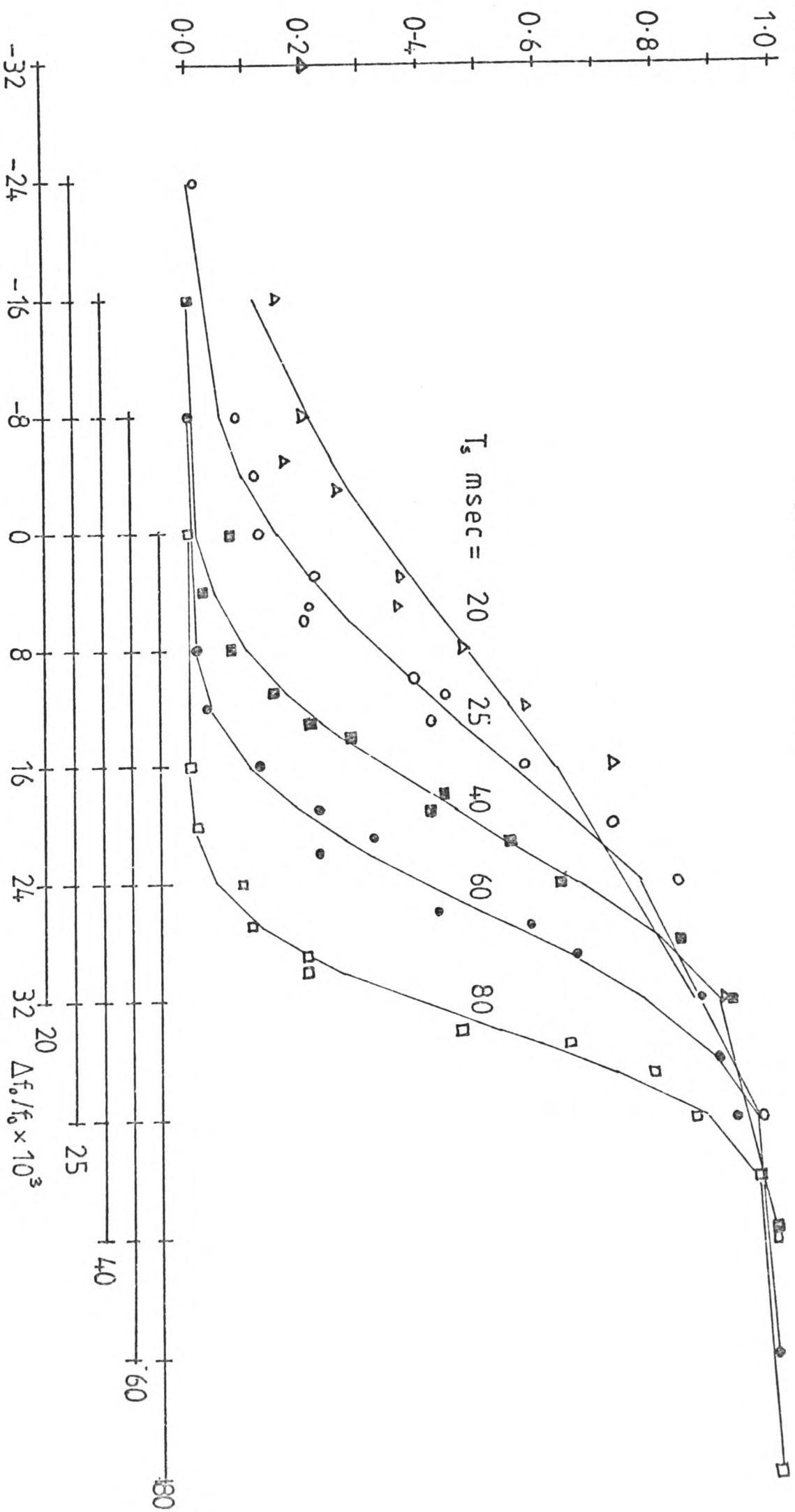
Symbol	T_s msec	Trials per point.	
		A.F.	G.S.
Δ	20	150	100
\circ	25	100	100
\blacksquare	40	125	75
\bullet	60	75	75
\square	80	75	75

The solid line in panels a) to d) is drawn through the best-fitting points of the theoretical function.

Panel e) shows the observed functions for 20 msec signals including data from the largest value of $\underline{\Delta f}/f$ of 0.06496.

$P(\text{"high"})$

Fig. 8.3a) O.A.F. $n=4$



$P(\text{"high"})$

Fig. 8.3 b) 0. A.F. $n=5$

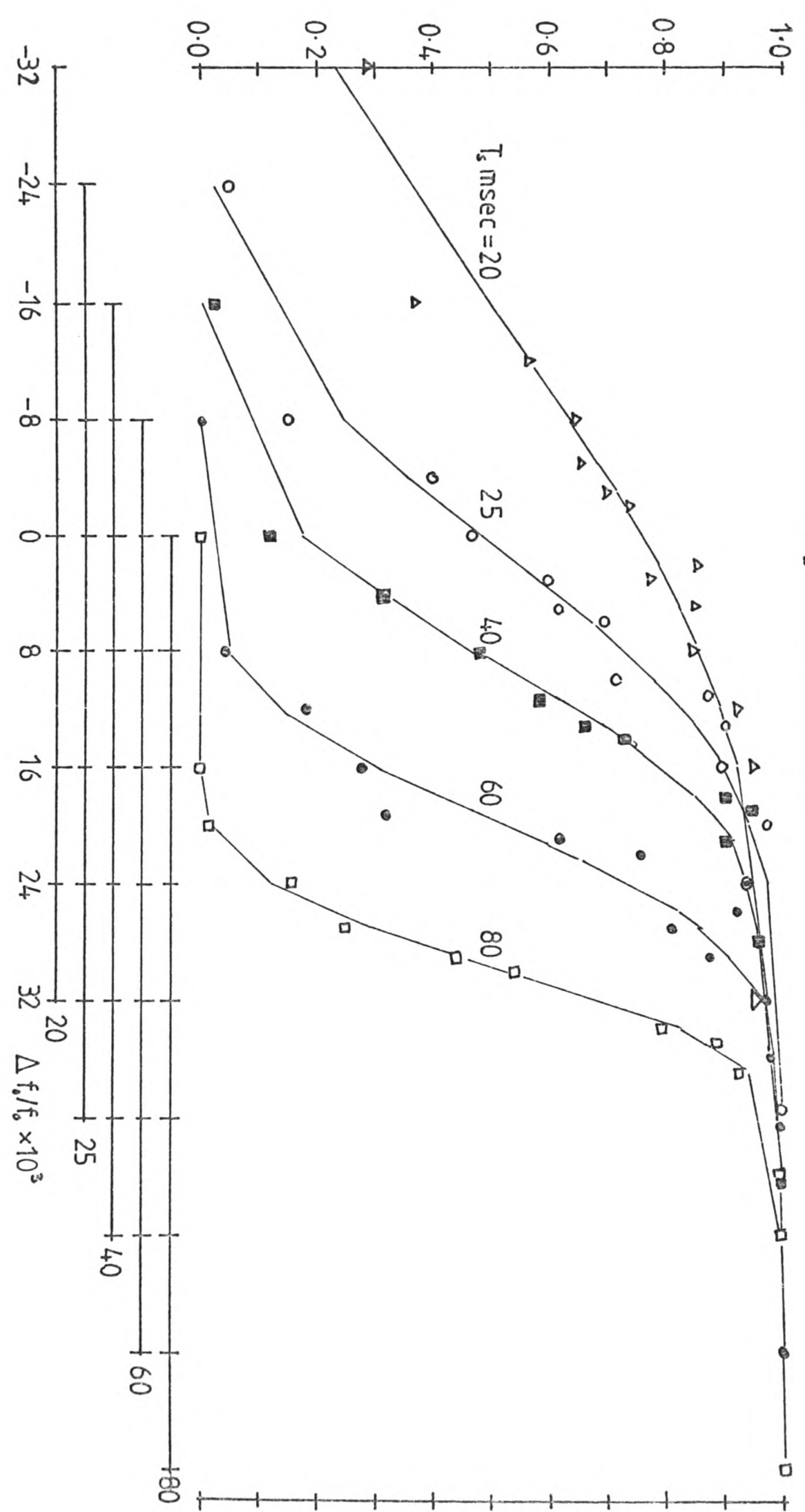
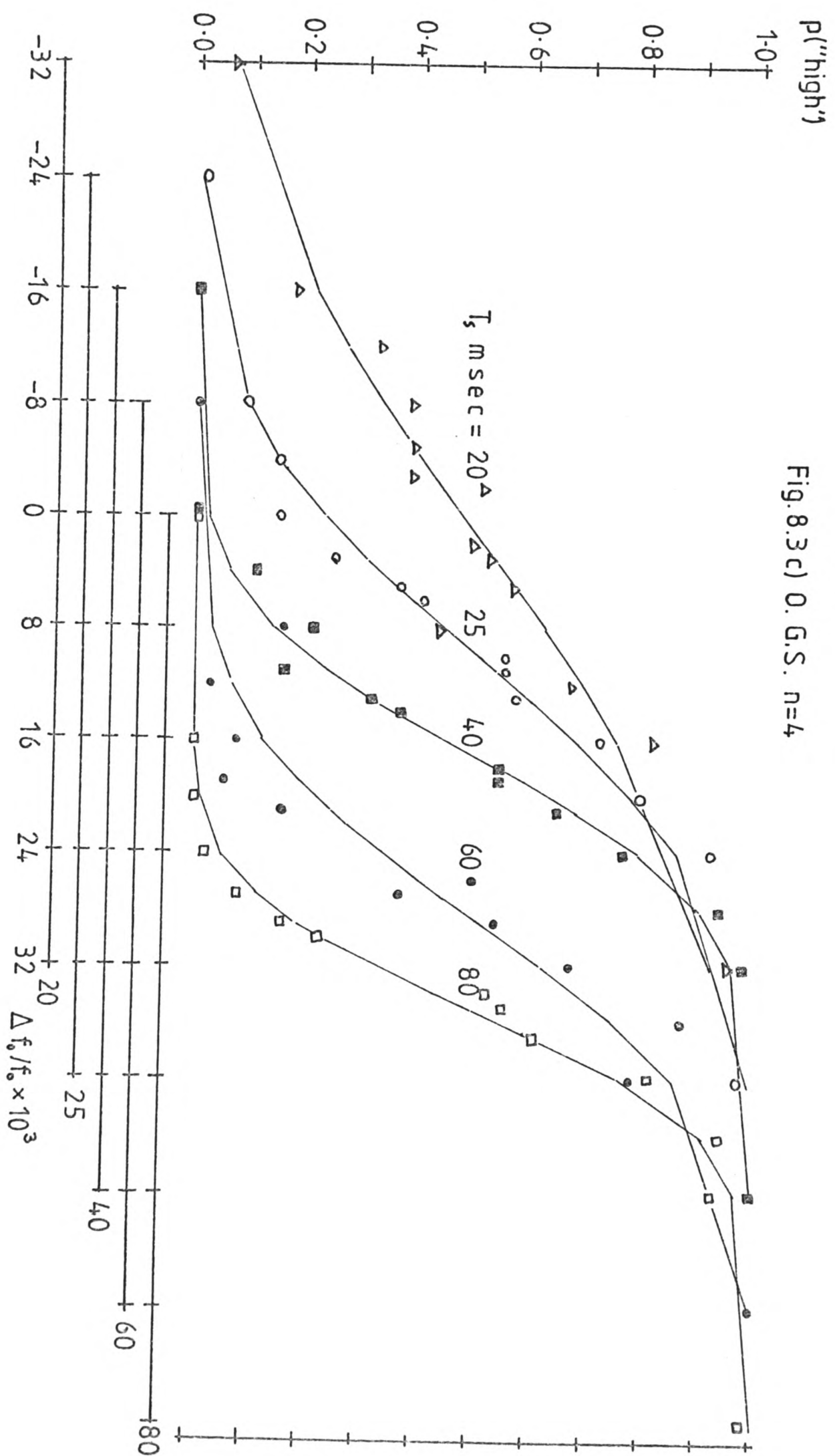


Fig. 8.3 c) 0. G.S. $n=4$



$P(\text{"high"})$

Fig. 8.3 d) O.G.S. $n=5$

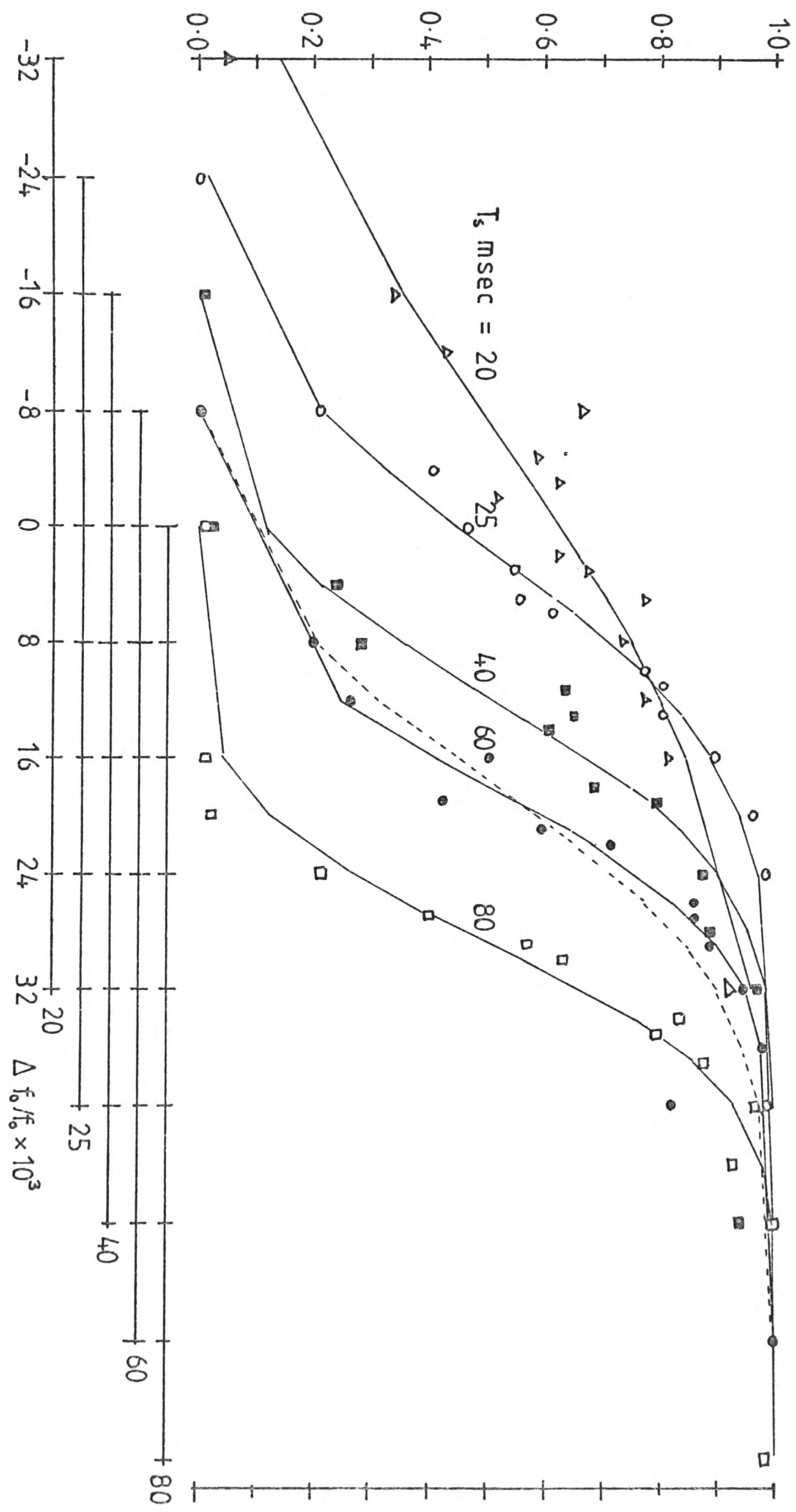


Fig. 8.3 e)

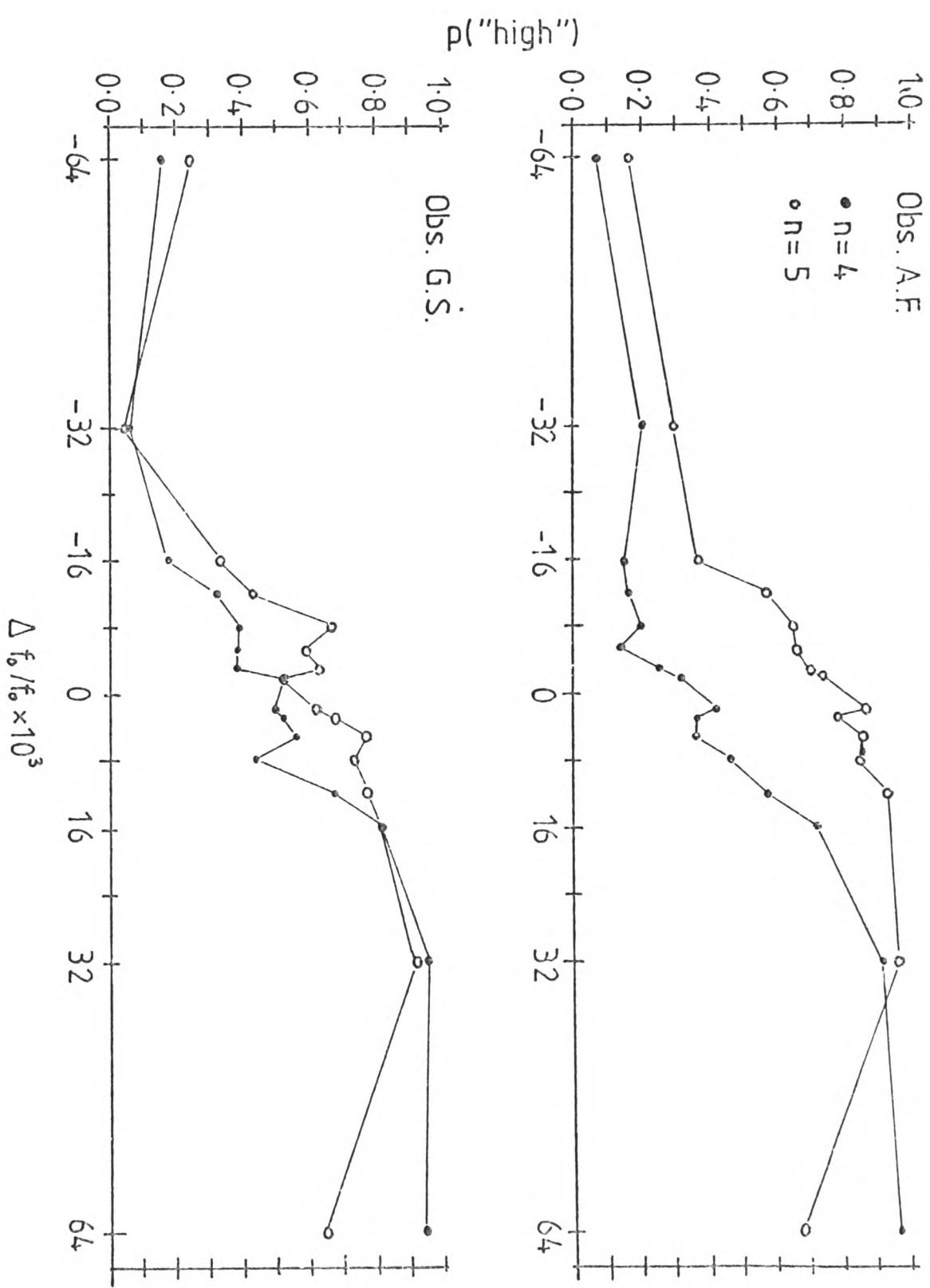
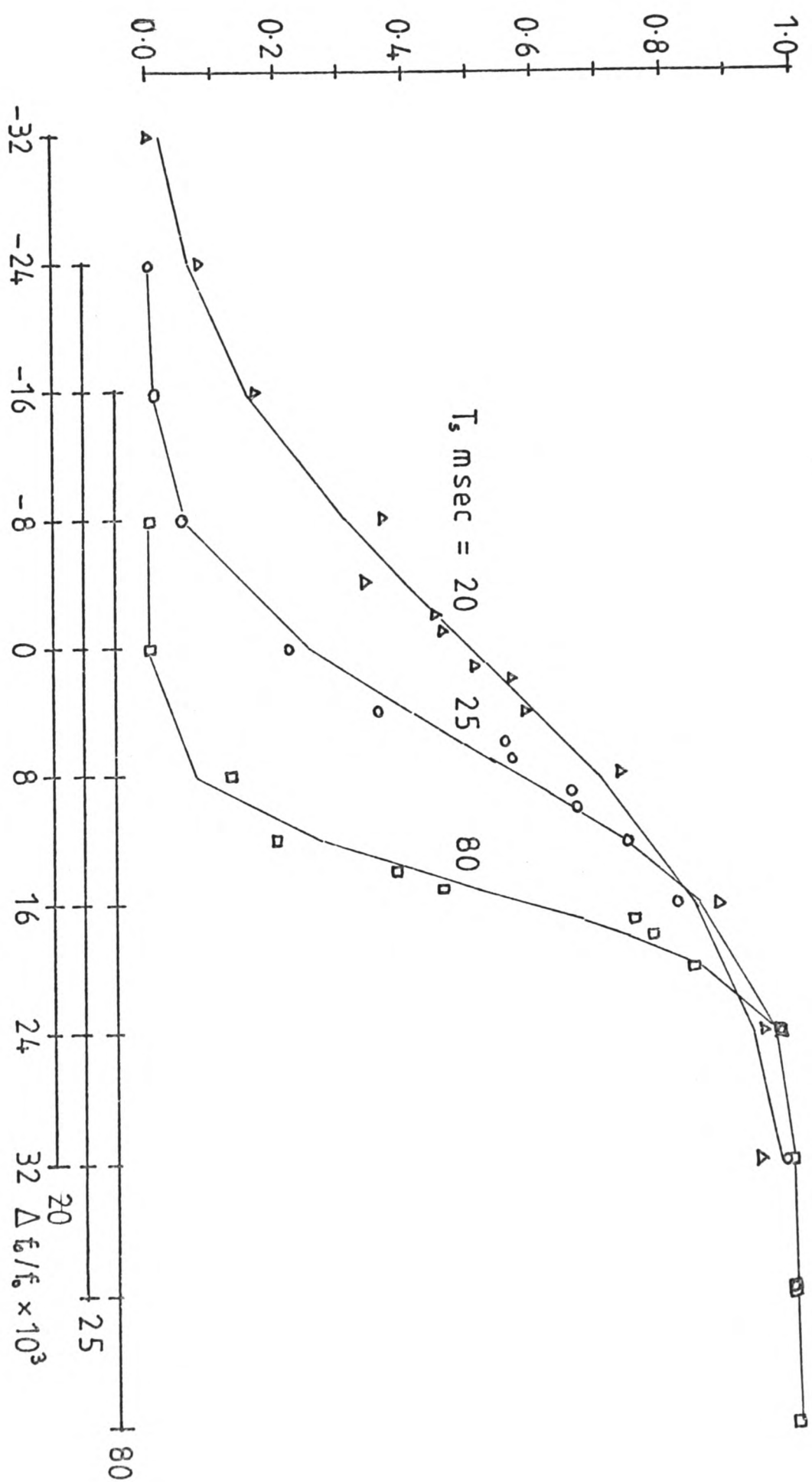


Figure 8.4. Psychometric functions from Experiment 13; pitch discrimination for a two-component complex signal.
Panel a), observer A.F., 100 trials per point.
Panel b), observer G.S., 100 trials per point at 80 msec, 125 trials per point at 25 and 20 msec.
The symbols are as those in figure 8,3.

$P(\text{"high"})$

Fig. 8.4a) O.A.F. $n=4$



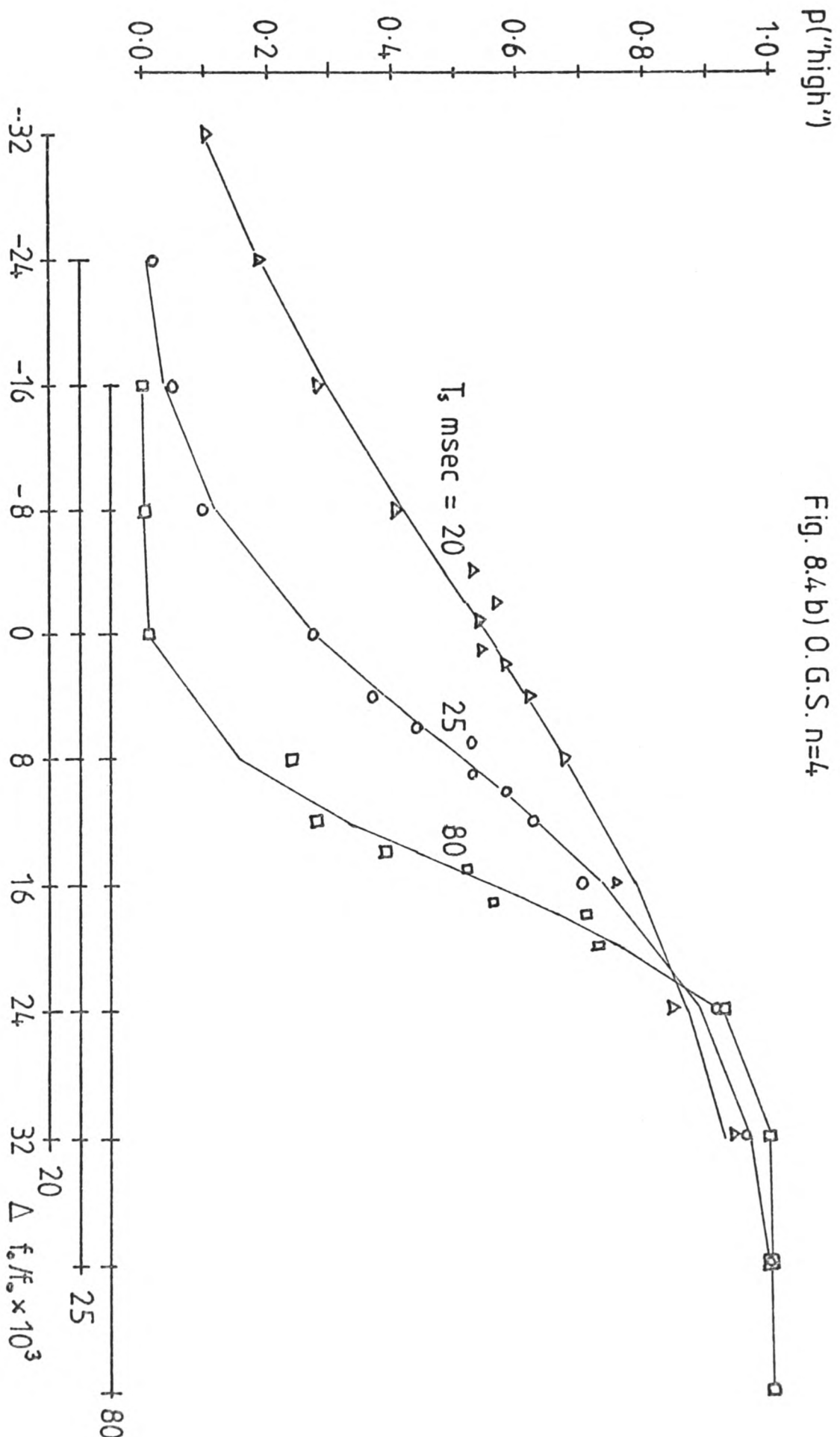


Table 8.1. Probit analysis of observed psychometric functions from Experiment 12 - n is the lower harmonic number of comparison signal.

Observer A.F.										
	T _s msec	N	Slope	S.E.	50% Point	S.E.	χ ²	df	hf	p
n=4	80	75	176.33	10.58	+0.0015	0.0003	10.013	11	1.0	<0.01
	60	75	125.31	7.41	+0.0016	0.0004	13.146	12	1.0	
	40	125	105.99	5.03	+0.0038	0.0004	8.439*	9	1.0	
	25	100	71.67	4.04	+0.0060	0.0006	12.747*	10	1.0	
	20	150	38.67	7.11	+0.0086	0.0027	115.354	12	9.613	
n=5	80	75	212.60	13.29	-0.0024	0.0003	4.165	11	1.0	
	60	75	140.68	8.22	-0.0042	0.0004	18.873	12	1.0	
	40	125	109.04	5.38	-0.0078	0.0004	12.856*	8	1.0	
	25	100	81.11	4.51	-0.0097	0.0019	12.160*	9	1.0	
	20	150	44.80	2.53	-0.0162	0.0010	19.566	11	1.0	
Observer G.S.										
	T _s msec	N	Slope	S.E.	50% Point	S.E.	χ ²	df	hf	p
n=4	80	75	147.06	8.71	+0.0030	0.0004	7.741*	6	1.0	<0.01
	60	75	129.21	9.10	+0.0039	0.0004	11.864	10	1.0	
	40	75	118.30	7.02	+0.0010	0.0004	14.199	12	1.0	
	25	100	75.14	4.18	+0.0016	0.0005	9.227*	10	1.0	
	20	100	46.37	4.41	+0.0010	0.0012	26.475	12	2.206	
n=5	80	75	130.09	7.75	-0.0033	0.0004	13.308	7	1.0	<0.01
	60	75	114.19	8.65	-0.0063	0.0006	9.608	10	1.0	
	40	75	102.13	11.90	-0.0047	0.0009	33.096	10	3.31	
	25	100	84.22	4.65	-0.0068	0.0005	7.383*	9	1.0	
	20	100	43.20	4.60	-0.0077	0.0015	30.093	12	2.508	

Table 8.2. Probit analysis of observed psychometric functions from Experiment 13 - n is the lower harmonic number of the two test signals.

Observer A.F.

	T _s msec	N	Slope	S.E.	50% Point	S.E.	χ ²	df	hf	p
n=4	80	100	211.84	14.03	-0.0012	0.0002	13.801	9	1.0	
	25	100	108.06	6.48	-0.0019	0.0004	8.040	12	1.0	
	20	100	63.43	3.37	-0.0003	0.0006	14.403	14	1.0	

Observer G.S.

	T _s msec	N	Slope	S.E.	50% Point	S.E.	χ ²	df	hf	p
n=4	80	100	142.91	9.18	-0.0010	0.0003	12.398	13	1.0	
	25	100	75.00	3.60	-0.0004	0.0005	6.531	14	1.0	
	20	100	41.11	2.06	-0.0033	0.0008	5.251	14	1.0	

Of the observed functions from Experiment 12, for observer G.S., seven of the ten functions showed no significant deviation from binomial variability, while for observer A.F., nine of the ten functions showed no such deviation. At signal

durations of 25 msec or more, none of the deviations were systematic, but for a signal duration of 20 msec, there was a tendency for performance at the maximum value of $|\Delta f|$ to be worse than that at the next smallest value. This effect was very apparent for $n=5$, and is shown in figure 8.1 e). The maximum value of $|\Delta f|$ was 0.06496×209.4 Hz (13.6 Hz), and the explanation of this effect appears to be that, particularly for the complex composed of the 5th and 6th harmonics, this frequency difference is approaching half of the frequency difference between the residue pitch mode at the true fundamental frequency and the adjacent lower frequency mode. The frequency differences between these two modes are 39 Hz for the complex of the 4th and 5th harmonics, and 33 Hz for that of the 5th and 6th harmonics. Because of this effect, the functions for a signal duration of 20 msec were reanalysed with the data from the maximum value of $|\Delta f|$ omitted. This removed the systematic deviation from the theoretical function; it is the results of these analyses which are included in the main panels of figure 8.3 and in table 8.1, and the parameters of these theoretical functions were those employed for the estimation of $\sigma(f'_0)/f_0$.

None of the ~~three~~ functions for each observer from Experiment 13 showed any significant deviation from the theoretical functions.

The Estimation of $\sigma(f'_0)/f_0$.

Experiment 12 employed a reference signal which differed

from the two-component comparison signal in structure as well as in frequency, so that $\frac{\sigma(f'_0)}{f_0}$ was estimated as $\frac{\sigma^2(x'_2 - x'_1)}{\sigma^2(x'_1)}$, where $\sigma(x'_2 - x'_1)$ was estimated as the reciprocal of the slope of the theoretical function, and $\sigma(x'_1)$ was taken as the relative DL for the reference signal as estimated in Experiment 4. The two signals in Experiment 13 differed only in frequency, and in this case, $\frac{\sigma(f'_0)}{f_0}$ was estimated as $\frac{\sigma(x'_2 - x'_1)}{\sqrt{2}}$. The estimates of $\frac{\sigma(f'_0)}{f_0}$ as a function of the signal duration for each of the two observers are shown in figure 8.5 together with their 95% confidence limits. Slopes corresponding to $1/\sqrt{T_3}$ and $1/T_3^{3/4}$ are also shown on the figure. It is apparent that, for signal durations of 25 msec and more, the former slope provides a good description of the data, while the slope between 25 and 20 msec is steeper. The signal duration dependence seen here is thus consistent with that shown in the studies reviewed in Chapter 3, and also consistent with both the Energy Detector model and the modification of Siebert's (1970) temporal model described by Goldstein and Sruлович (1977).

A Comparison of the Effects of Duration and Signal-to-Noise Ratio.

The trading relationship between signal duration and signal-to-noise ratio predicted by the Energy Detector model can be examined by a comparison of the estimates of $\frac{\sigma(f'_0)}{f_0}$ from Experiments 8 and 12, and of those from Experiments 10 and 13. Figures 8.6 and 8.7 display the estimates from the comparable experiments, where signal duration is scaled in terms of the

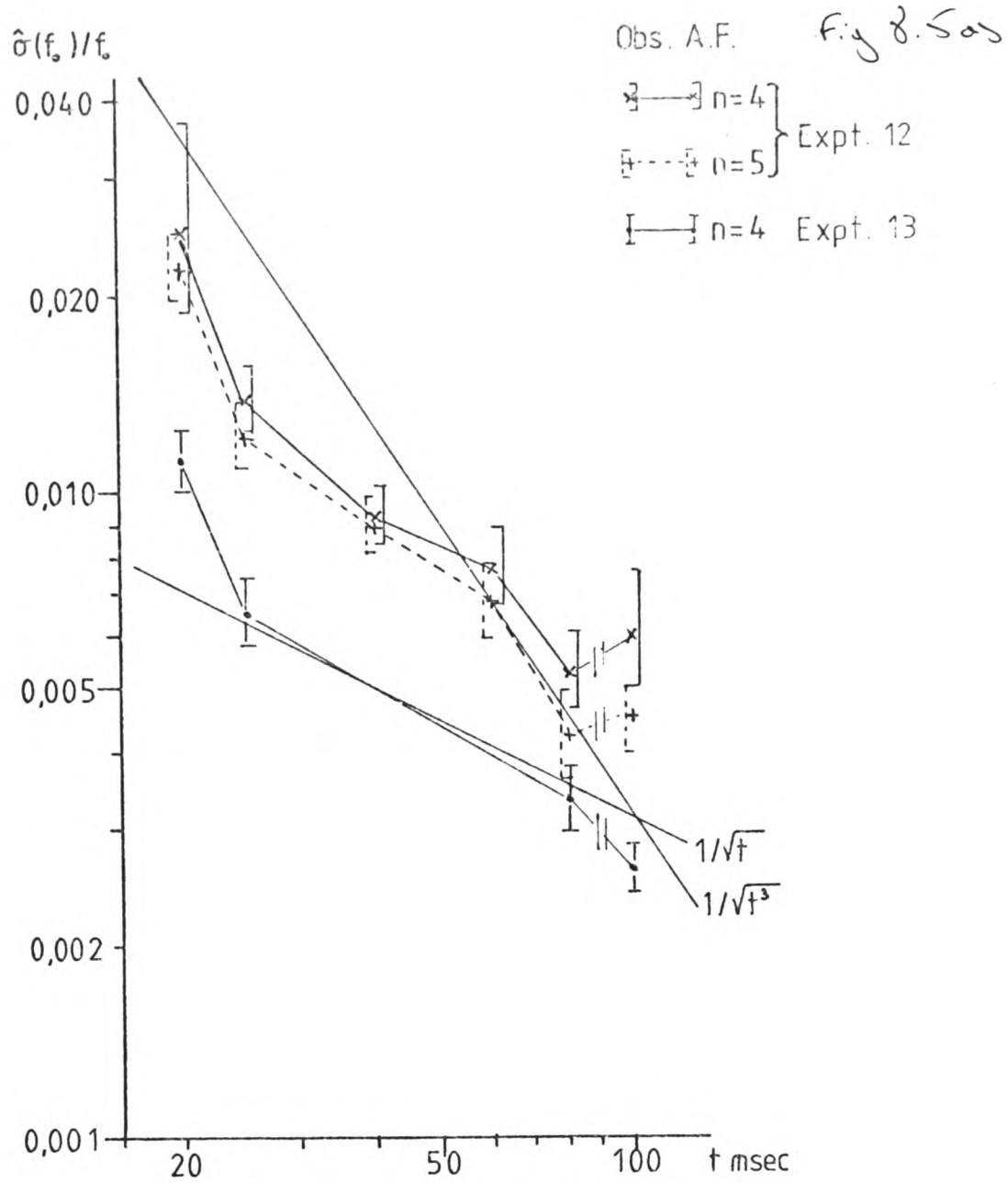


Fig. 8.5 b)

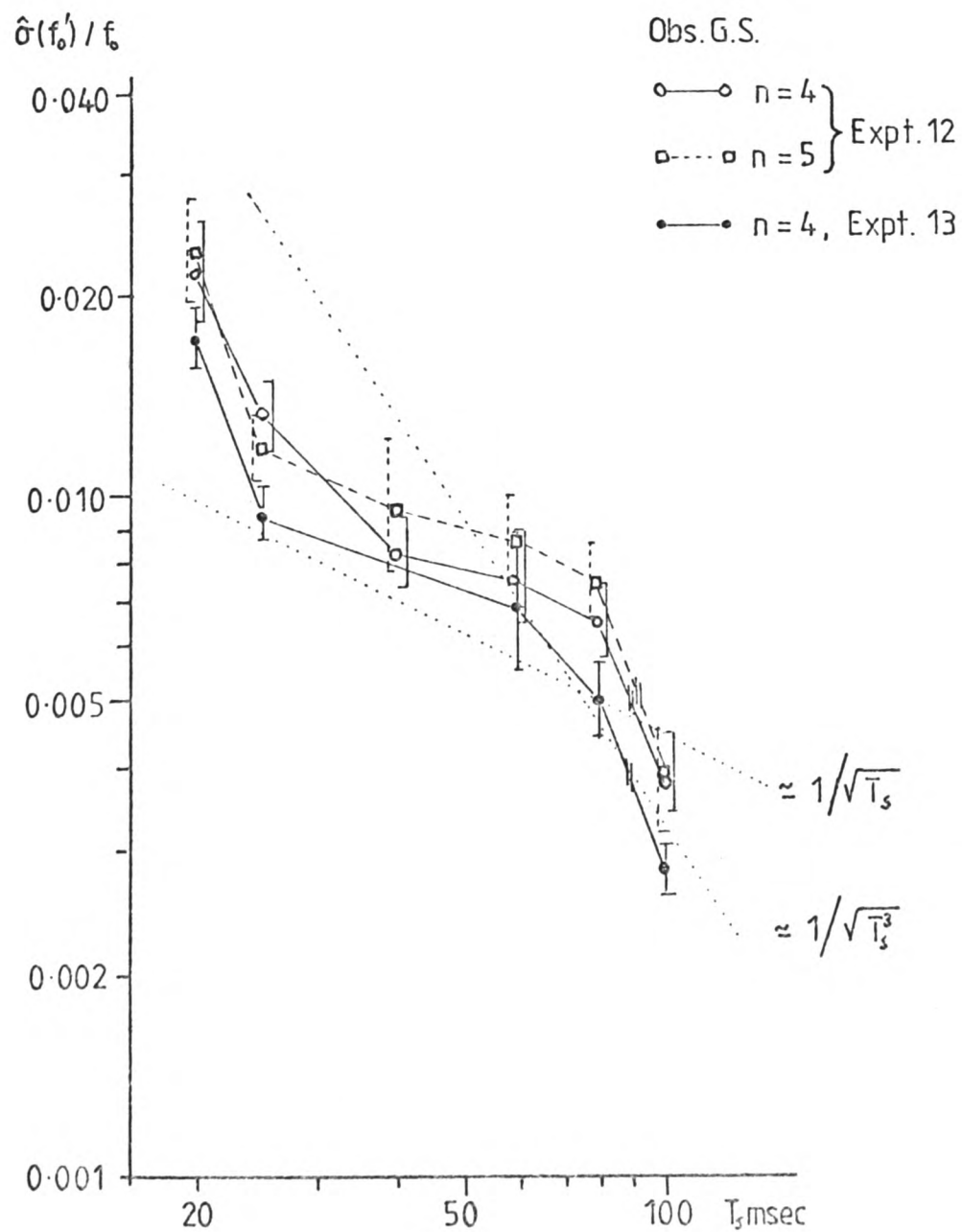


Figure 8.5. Estimates of $\hat{\sigma}(f'_0)/f_0$ as a function of signal duration. Data from Experiments 12 and 13, except for the points at 100 msec, which are from Experiments 8 and 10, for $E-N_0$ of 30 dB. The error bars are 95% confidence limits. Also shown are slopes of $1/\sqrt{T_s}$ and $1/\sqrt{T_s^3}$.

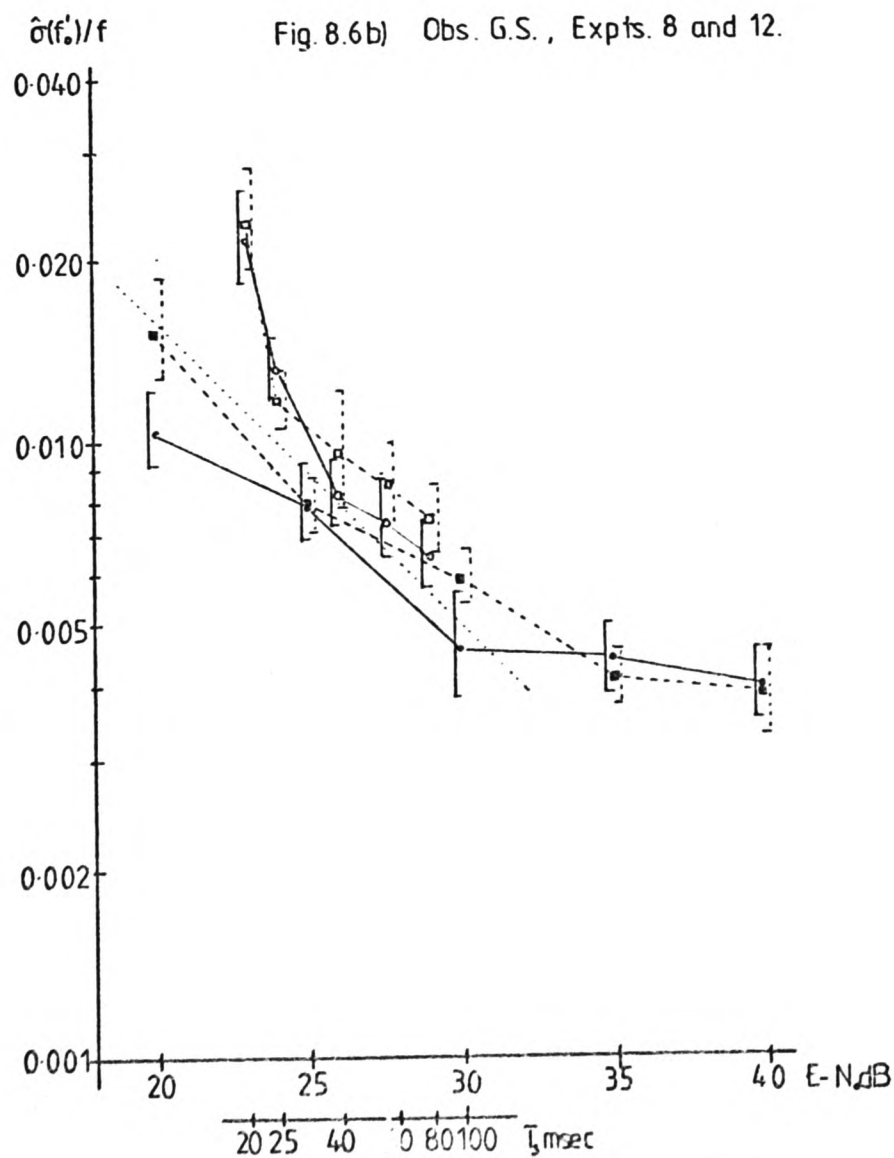
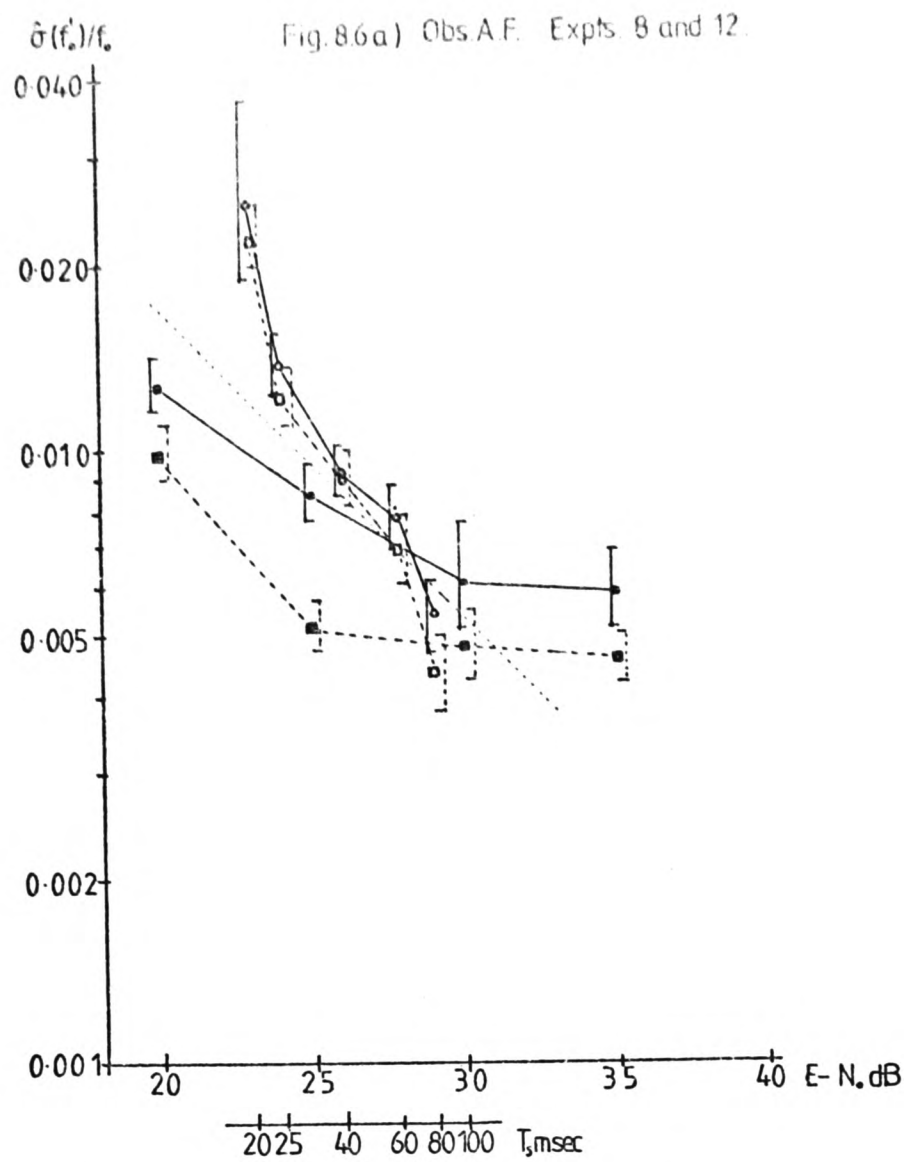


Figure 8.6. A comparison of the estimates of $\sigma(f'_0)/f_0$ for equivalent signal energy to noise energy ratios. Data from Experiments 8 and 12. Panel a), observer A.F., panel b), observer G.S. The filled symbols are for signals of constant duration presented with a variable noise level, and the empty symbols are for signals of variable duration presented with a constant noise level. The solid lines and circular symbols refer to the two-component signal having the lower harmonic number n of 4, and the square symbols the signal having $n = 5$. The dotted line has a slope of $1/\sqrt{N_0}$, which is equivalent to a slope of $1/\sqrt{T_0}$. The error bars are 95% confidence limits.

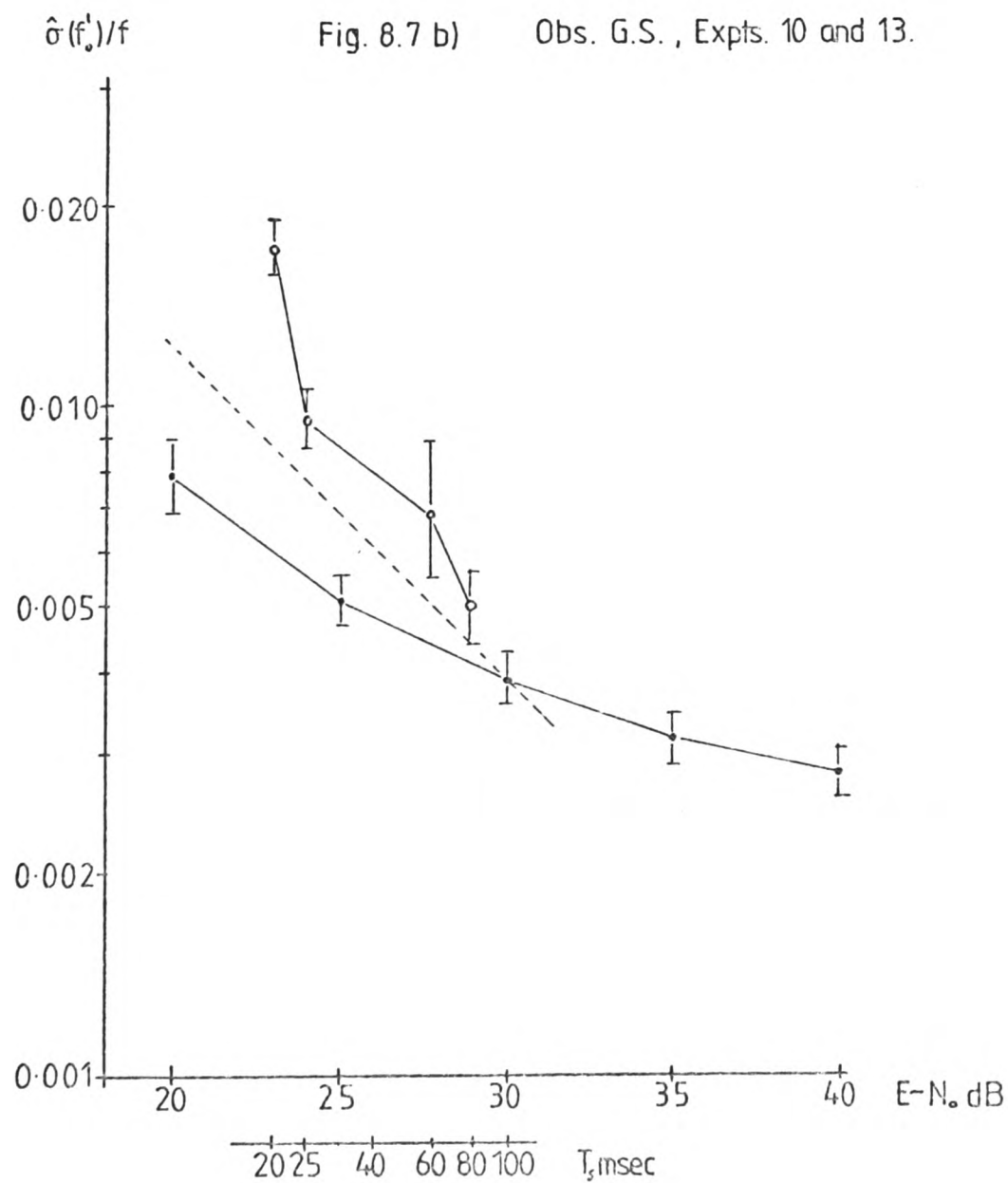
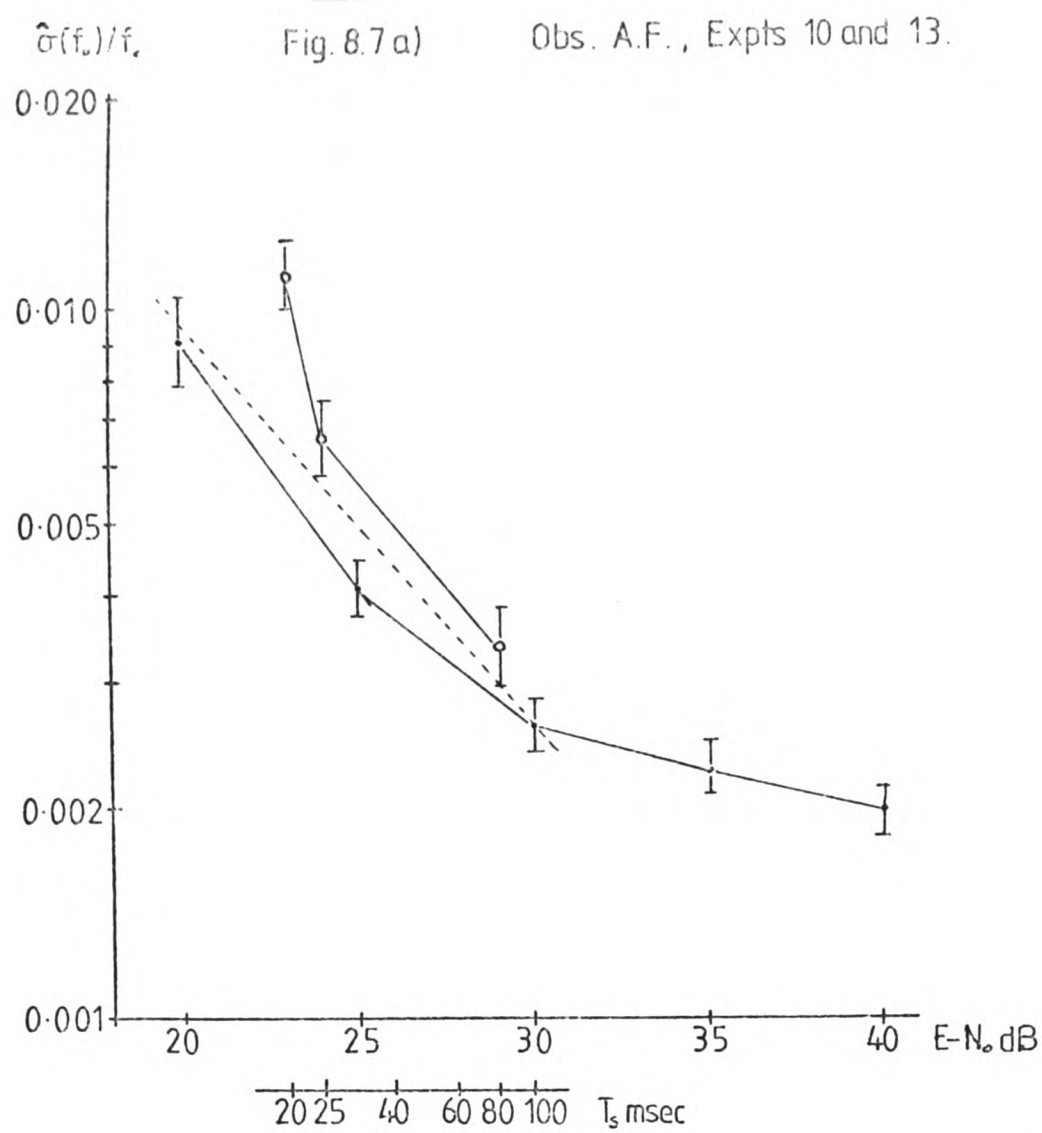


Figure 8.7. A comparison of estimates of $\sigma(f'_0)/f_0$ at equivalent signal-energy to noise energy ratios. Data from Experiments 10 and 13. Panel a), observer A.F., panel b), observer G.S. Filled symbols are for signals of constant duration, presented against a variable noise level. Empty symbols are for signals of variable duration, with a constant noise level. The dashed line has a slope of $1/\sqrt{N_0}$, which is equivalent to a slope of $1/\sqrt{T_s}$. The error bars are 95% confidence intervals.

signal energy to noise energy ratio in dB. The signal energy was calculated relative to that of a 100 msec signal. At equivalent signal-to-noise ratios, the estimates of $\sigma(f'_0)/f_0$ for the shorter duration signals are increasingly larger as signal duration decreases, and for signal durations of less than 80 msec, these differences are greater than the 95% confidence limits for the estimates. Therefore, the predictions of the Energy Detector model are not entirely consistent with the data. With the exception of the estimate of $\sigma(f'_0)/f_0$ at 20 msec, the dependence of the estimates upon duration is consistent with a dependence on $1/\sqrt{T_s}$, so these data in themselves are in accordance with the predictions of the Energy Detector model. There may be an intercept difference between these sets of data, which could well be due to the difference in the rise and fall times of the signals, which were shorter for the variable duration signals (Ronken, 1971). The estimates of $\sigma(f'_0)/f_0$ as a function of the background noise level for observer A.F., and for observer G.S. in Experiment 8, are consistent with a dependence upon $\frac{1}{\sqrt{N_0}}$, at least for $E-N_0$ of 30 dB and less, so that for these three comparisons, an intercept difference could largely account for the differences found with the two manipulations.

This simple explanation cannot be true for the performance of observer G.S. in Experiments 10 and 13, where the observed dependence of $\sigma(f'_0)/f_0$ on noise level has a slope smaller than $\frac{1}{\sqrt{N_0}}$. This noise level dependance was, however, reasonably well-fitted by the Energy Detector model (see Chapter 7). Slope differences of this sort would also be predicted by a modified

Energy Detector model including an additive internal noise whose magnitude is a function of signal power. With such a modification, the dependence of discrimination on external noise will be show a slope of less than $1/\sqrt{N_0}$, while the dependence on duration will be unaffected by the internal noise.

The Differences Between Discrimination Performance in the Two Tasks.

The results reported here also bear upon an issue considered in the previous two chapters, this being the difference between pitch discrimination performance for tasks involving comparisons between and within frequency-selective channels. As was the case in previous experiments, the estimates of $\sigma(f'_0)/f_0$ from the present two experiments are larger where pitch comparisons must be made between such channels. Considering only the signals containing the 4th and 5th harmonics, the estimates of $\sigma(f'_0)/f_0$ from Experiment 12 exceed those from Experiment 13 by an average factor of 1.857 and 1.262 for observers A.F. and G.S. respectively, and these factors are similar to those reported above. The factor for observer G.S. is close to the factor of 1.22 expected from the difference in stimulus uncertainty between the two tasks, but the factor for observer A.F. is significantly larger than that value, and increases with decreasing duration. A possible explanation of this effect is discussed in the following chapter.

Summary.

Discrimination for the residue pitch of a two-component harmonic signal with a fundamental frequency of about 200 Hz was measured as a function of signal duration.

Conclusions.

A/ The observed duration dependence of the estimates of $\sigma(f'_0)/f_0$ was consistent with an inverse dependence upon the square root of signal duration for durations between 25 and 80 msec, while the increase in the estimates of $\sigma(f'_0)/f_0$ between durations of 25 and 20 msec was greater. The observed dependence was thus consistent with an Energy Detector model of frequency discrimination (Green and Swets, 1966; Henning 1967a), and also with the temporal model described by Goldstein and Sruлович (1977).

B/ The obtained estimates of $\sigma(f'_0)/f_0$ were larger than those obtained at equivalent signal-to-noise energy ratios where signal power and duration were constant, and noise energy was varied, the differences increasing with decreasing duration. This result is not consistent with an unmodified Energy Detector model, but can be interpreted in terms of such a model if it is assumed that an internal noise is present. A difference in the rise and fall times of the signals may also have contributed to these differences.

C/ Pitch discrimination performance in a task where the

compared pitches are derived from components in different frequency regions was found to be inferior to performance in a task where the signal components are of similar frequency. This result is consistent with the results of the experiments reported in the previous two chapters.

CHAPTER 9

Abstract.

This chapter reviews the results of the pitch and frequency discrimination experiments described in the previous four chapters. These results are discussed primarily in terms of Goldstein's (1973) spectral-pattern recognition theory of residue pitch. Goldstein's assumption that the representation of component frequency, as seen by the pitch processor, includes added internal noise is examined in detail, and estimates of the extent of such an internal noise are made from the pitch discrimination data. It is concluded that pitch comparisons which may be made by reference to multiple within-channel comparisons of frequency differences are unaffected by such an internal noise, and that the variability of such judgements is due only to the variability of the peripheral coding of frequency, and to external noise. Where pitch judgements are made between two signals affecting two different sets of frequency selective channels, an internal noise would lead to a performance decrement, in agreement with the results described

above. A modification of Goldstein's model, allowing the sign of a pitch difference to be estimated from within-channel frequency differences before the pitch itself is extracted, allows the model to provide an excellent account of the performance of one observer, and a fair account of the performance of a second. A comparable modification to Wightman's (1973b) autocorrelation model would have the same consequences.

The value of the postulated internal noise was estimated from performance in experiments requiring both within and between-channel comparisons. On a scale of relative frequency, the internal noise variance was estimated as about 10^{-5} . The corresponding standard deviation is about three times smaller than the value of 0.01 assumed by Goldstein (1973) in describing performance in musical interval identification. Goldstein's estimation of this parameter would however include any external noise, the noise due to peripheral coding, and also an effect of stimulus uncertainty.

The Predictions of Residue Pitch Theories and the Two Fundamental Frequency Discrimination Tasks.

The generalised prediction made in Chapter 2 from the spectral-pattern recognition theories of residue pitch was that $\sigma(f'_0)/f_0$, the relative variance of the mode of the pitch pdf having its mean around the true fundamental frequency f_0 of a two-component harmonic complex, is determined by the reciprocal sum of the relative variances $(\sigma(f'_j)/f_j)^2$, where these variances refer to the internal representations f'_j of the component frequencies f_j of the complex, and hence, $\sigma(f'_0)/f_0$ should be smaller than the minimum of $\sigma(f'_j)/f_j$. The parameter $\sigma(f'_j)/f_j$ was estimated for the 4th and 5th harmonics of 200 Hz as a function of signal level in Experiment 3, and as a function of $E-N_0$ in Experiment 7. Estimates of $\sigma(f'_0)/f_0$ from Experiments 6 and 10, each of which involved a comparison of the residue pitches of a pair of complex signals containing the 4th and 5th harmonics of a fundamental frequency close to 200 Hz, were almost completely consistent with the predictions based on the estimates of $\sigma(f'_j)/f_j$ from Experiments 3 and 7. The periodicity theory, which predicts that $\sigma(f'_0)/f_0$ cannot be smaller than the minimum of $\sigma(f'_j)/f_j$, could be clearly rejected for the comparison of Experiments 3 and 6, but not for Experiments 7 and 10, where the estimates of $\sigma(f'_j)/f_j$ for the two components differed by a factor of about 1.5, so that the minimum of these estimates was

approximately equal to the prediction from the spectral-pattern recognition theories.

However, where $\sigma(f'_0)/f_0$ was estimated from tasks requiring the comparison of two signals with no common components, the resulting estimates always exceeded the values predicted from the variability of the component frequency estimates. Some part of this difference may be attributed to the effects of stimulus uncertainty, at least for observer A.F., for whom data are available from Experiments 8 and 9, which differed only in the uncertainty of the harmonics of the two-component comparison signal. However, such an explanation does not completely account for the discrepancy between the obtained estimates of $\sigma(f'_0)/f_0$ from tasks requiring comparisons between pitches derived from components in the same and in different frequency regions. This discrepancy may indicate that where the latter comparisons are made, pitch is derived from periodicity information rather than from independent estimates of the component frequencies. This explanation, however, is not consistent with the results of pitch comparisons of the former kind, where, it is clear that the judgements were based upon a joint perception of the two components, and that the available periodicity information cannot account for the observed performance. For reasons of parsimony, it seems undesirable to invoke a different perceptual mechanism to account for performance in the latter tasks.

A summary of the estimates of $\frac{\sigma(f'_j)}{f_j}$ and $\frac{\sigma(f'_0)}{f_0}$ from Experiments 3 to 13 is presented as table 9.1.

Table 9.1 Estimates of $\sigma(f'_j)/f_j$ and $\sigma(f'_0)/f_0$ and their 95% confidence limits as a function of signal level. The values shown are $\sigma(f')/f \times 10^3$. The legends 'same' and 'diff't.' refer to the channels employed for the pitch comparison. The legend 'n' refers to the harmonic number of the test signal for Experiments 3 and 7, and to the lower harmonic of the two-component signal for the other experiments.

Estimates as a function of sensation-level.						
Obs Expt.	SL dB	15	20	30	40	50
A.F. 3	$\sigma(f'_j)/f_j$	7.559	3.499	2.259	2.273	3.356
n=4	upper 95% limit	8.356	3.750	2.451	2.465	3.727
	lower 95% limit	6.901	3.280	2.095	2.108	3.052
A.F. 3	$\sigma(f'_j)/f_j$	4.472	2.634	2.357	2.006	3.487
n=5	upper 95% limit	4.943	2.854	2.642	2.181	3.868
	lower 95% limit	4.033	2.446	1.984	1.855	3.174
	predicted $\sigma(f'_0)/f_0$	3.849	2.104	1.554	1.504	2.418
A.F. 5	$\sigma(f'_0)/f_0$	4.388	4.438	3.415	3.699	3.859
n=4	upper 95% limit	5.150	5.724	4.277	4.429	4.369
diff't	lower 95% limit	3.785	3.046	2.711	3.117	3.431
A.F. 5	$\sigma(f'_0)/f_0$	4.794	5.565	4.179	3.172	4.559
n=5	upper 95% limit	5.944	7.348	4.725	3.528	5.819
diff't	lower 95% limit	3.966	4.413	3.722	2.861	3.683
A.F. 6	$\sigma(f'_j)/f_j$	2.724	2.307	2.001	2.025	2.350
n=4	upper 95% limit	2.989	2.521	2.191	2.215	2.579
same	lower 95% limit	2.501	2.126	1.841	1.865	2.174
Obs Expt.	SL dB	15	20	30	40	50
G.S. 3	$\sigma(f'_j)/f_j$	3.719	3.054	2.783	3.228	4.510
n=4	upper 95% limit	4.309	3.278	2.983	3.451	4.984
	lower 95% limit	3.270	2.858	2.608	3.029	4.118
G.S. 3	$\sigma(f'_j)/f_j$	4.976	3.088	2.923	3.830	6.634
n=5	upper 95% limit	5.495	3.356	3.171	4.117	7.624
	lower 95% limit	4.546	2.860	2.711	3.580	6.063
	predicted $\sigma(f'_0)/f_0$	2.979	2.171	2.016	2.468	3.730
G.S. 5	$\sigma(f'_0)/f_0$	4.318	3.490	2.742	4.320	5.864
n=4	upper 95% limit	4.805	4.085	3.110	4.806	6.539
diff't	lower 95% limit	3.899	2.999	2.415	3.902	5.297
G.S. 5	$\sigma(f'_0)/f_0$	4.586	4.022	3.570	4.293	6.185
n=5	upper 95% limit	5.124	4.495	4.123	4.805	7.088
diff't	lower 95% limit	4.091	3.613	3.108	3.895	5.461
G.S. 6	$\sigma(f'_0)/f_0$	2.799	2.521	2.177	2.396	3.731
n=4	upper 95% limit	3.100	2.712	2.377	2.574	4.017
same	lower 95% limit	2.548	2.355	2.007	2.241	3.484

Table 9.1 continued.

Estimates as a function of signal-to-noise ratio.

Obs	Expt.	E-N ₀	dB	20	25	30	35	40
A.F.	7	$\sigma(f'_j)/f_j$		12.039	6.490	4.396	4.363	4.142
	n=4	upper 95% limit		14.016	7.332	4.908	4.880	4.627
		lower 95% limit		10.557	5.822	3.974	3.945	3.749
A.F.	7	$\sigma(f'_j)/f_j$		8.597	4.223	2.908	2.538	2.936
	n=5	upper 95% limit		9.769	4.720	3.221	2.863	3.254
		lower 95% limit		7.767	3.821	2.650	2.361	2.675
		predicted $\sigma(f'_0)/f_0$		6.996	3.532	2.425	2.223	2.406
A.F.	10	$\sigma(f'_0)/f_0$		8.877	4.042	2.608	2.275	1.971
	n=4	upper 95% limit		10.300	4.435	2.848	2.484	2.157
		same lower 95% limit		7.799	3.712	2.406	2.099	1.815
A.F.	8	$\sigma(f'_0)/f_0$		12.634	8.477	6.163	5.915	-
	n=4	upper 95% limit		14.076	9.422	7.657	6.917	-
		diff't lower 95% limit		11.537	7.694	5.116	5.142	-
A.F.	8	$\sigma(f'_0)/f_0$		9.891	5.192	4.822	4.594	-
	n=5	upper 95% limit		10.961	5.745	5.533	5.078	-
		diff't lower 95% limit		8.999	4.723	4.249	4.188	-
A.F.	9	$\sigma(f'_0)/f_0$		11.565	5.252	4.507	3.145	3.285
	n=4	upper 95% limit		13.172	5.929	5.087	3.582	4.004
		diff't lower 95% limit		10.293	4.696	4.026	2.774	2.725
Obs	Expt.	E-N ₀	dB	20	25	30	35	40
G.S.	10	$\sigma(f'_0)/f_0$		7.828	5.058	3.838	3.150	2.776
	n=4	upper 95% limit		8.950	5.489	4.207	3.443	3.031
		same lower 95% limit		6.956	4.639	3.529	2.903	2.561
G.S.	8	$\sigma(f'_0)/f_0$		10.378	7.869	4.640	4.373	3.940
	n=4	upper 95% limit		12.177	9.066	5.699	5.019	4.470
		diff't lower 95% limit		9.022	6.929	3.351	3.840	3.491
G.S.	8	$\sigma(f'_0)/f_0$		15.121	7.843	5.957	4.054	3.827
	n=5	upper 95% limit		18.561	8.724	6.595	4.512	4.544
		diff't lower 95% limit		12.736	7.110	5.417	3.658	3.253

Table 9.1 continued.
Estimates of $\sigma(f'_0)/f_0$ as a function of signal duration.

Obs	Expt. T, msec	20	25	40	60	80
	Equivalent E-No	23	24	26	27.8	29
A.F. 13	$\sigma(f'_0)/f_0$	11.147	6.544	-	-	3.338
	n=4 upper 95% limit	12.442	7.414	-	-	3.836
	same lower 95% limit	10.097	5.856	-	-	2.954
A.F. 12	$\sigma(f'_0)/f_0$	25.144	13.836	9.259	7.771	5.373
	n=4 upper 95% limit	36.991	15.580	10.204	8.842	6.166
	diff't lower 95% limit	19.020	12.436	8.439	6.918	4.740
A.F. 12	$\sigma(f'_0)/f_0$	22.250	12.195	8.990	6.874	4.341
	n=5 upper 95% limit	25.042	13.718	9.989	7.821	5.044
	diff't lower 95% limit	20.014	10.969	8.164	6.113	3.777
G.S. 13	$\sigma(f'_0)/f_0$	17.200	9.428	-	6.799	4.948
	n=4 upper 95% limit	19.074	10.407	-	8.830	5.661
	same lower 95% limit	15.661	8.618	-	5.528	4.394
G.S. 12	$\sigma(f'_0)/f_0$	21.466	13.146	8.195	7.457	6.477
	n=4 upper 95% limit	25.887	14.794	9.338	8.737	7.409
	diff't lower 95% limit	18.322	11.818	7.283	6.478	5.730
G.S. 12	$\sigma(f'_0)/f_0$	23.055	11.691	9.569	8.508	7.402
	n=5 upper 95% limit	28.493	13.151	12.239	10.072	8.453
	diff't lower 95% limit	19.346	10.513	7.814	7.339	6.565

A Modification to Goldstein's Model.

One assumption which can reconcile the results of the tasks requiring a comparison of the pitches of two signals having their components in different frequency regions with the spectral-pattern recognition models and the estimates of $\sigma(f'_j)/f_j$ from Experiments 3 and 7, is that some internal noise process contributes to the variance of either f'_j or f'_b at the stage of perceptual processing at which pitches derived from components of widely differing frequencies are compared. Such a noise may result, as Goldstein (1973) has proposed, from noisy channels conveying component frequency estimates to the pitch processor. Alternatively, this noise may be a component of the variability of the decision criterion, which is present only when such comparisons are made, and absent where comparisons may be made between pitches derived from frequency components coded in the same sensory channels. These explanations cannot be distinguished on the basis of the data described above. A modified version of Goldstein's model, which can account for the results described above, is illustrated in figure 9.1. This scheme differs from that presented by Goldstein in that each frequency selective channel incorporates a frequency comparator, which is able to code frequency differences, or at least the sign of a frequency difference, before the component frequency estimates are conveyed to the pitch processor by the noisy channels. A comparable modification may be made to the other

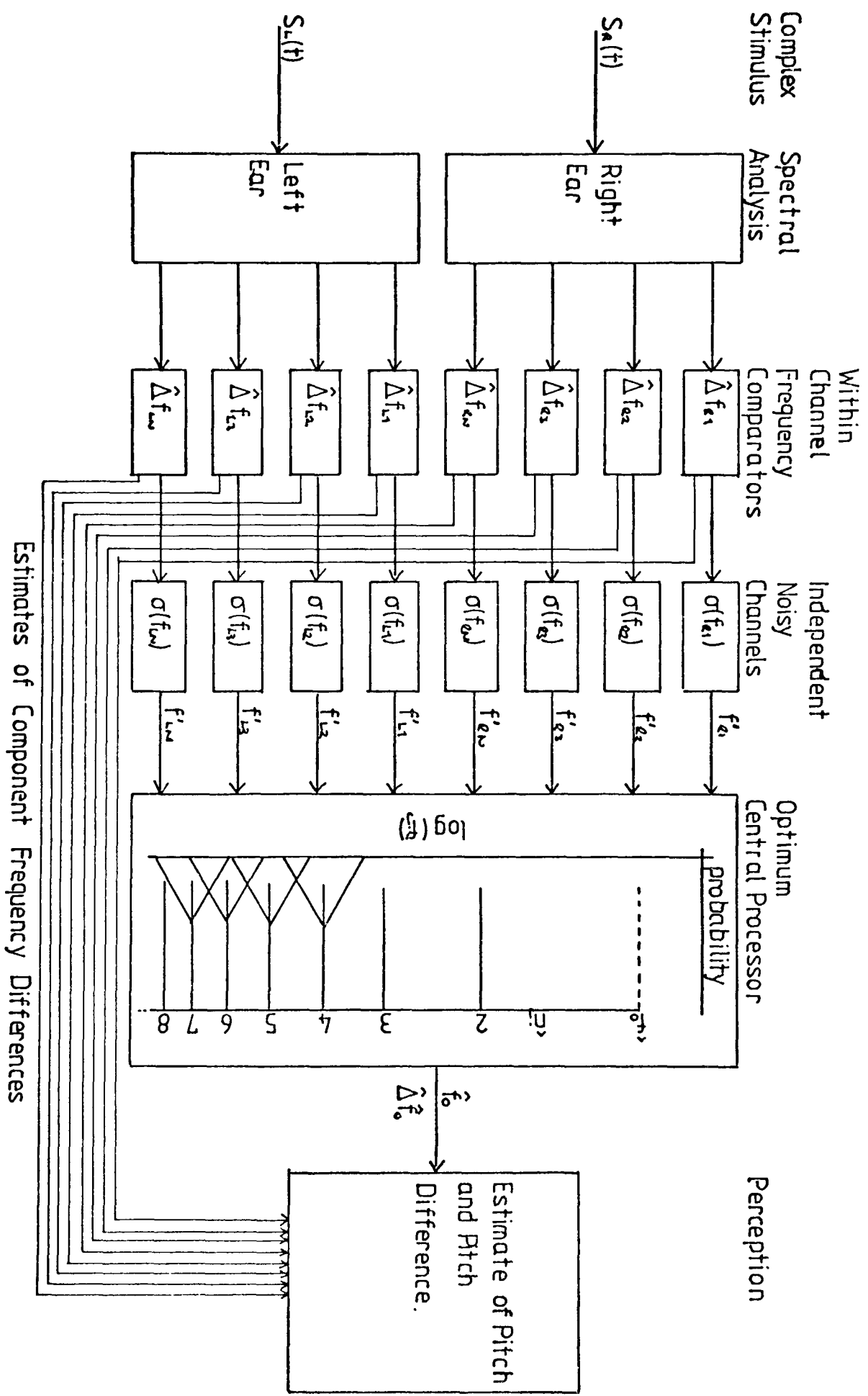


Figure 9.1. The proposed modification to Goldstein's (1973) optimum-processor model (shown unmodified in figure 2.1). The addition to the model is the series of within-channel frequency comparators.

spectral-pattern recognition models described earlier.

If it is assumed that an observer is able to restrict his sampling to those frequency selective channels which are likely to convey component frequency estimates, then the estimates of $\sigma(f'_0)/f_0$ from the various tasks described above allow an estimate of the noisiness added between the within channel frequency comparators and the between channel pitch comparators. Denoting estimates of $\sigma(f'_0)/f_0$ from a task requiring only within channel comparisons as $\hat{\sigma}_s$, and estimates from a task requiring between channel comparisons as $\hat{\sigma}_p$, this additional noise, here denoted σ_p , can be estimated as,

$$\hat{\sigma}_p^2 = \hat{\sigma}_b^2 - u \cdot \hat{\sigma}_s^2,$$

where u represents the uncertainty factor, that is, the number of channels contributing to σ_b divided by the number contributing to σ_s . If the noise is due to individual noisy channels rather than to a criterion variance, then the noise in each channel can be estimated as $\hat{\sigma}_p^2 \times (2/u)$: The term $1/u$ eliminates the contribution of frequency uncertainty, and if the internal noise is due to equal contributions from two channels which combine as a reciprocal sum, the contribution from each channel will be equal to twice the internal contribution to σ_p^2 . If the assumption of such an internal noise is correct, the estimates of σ_p should be independent of $\hat{\sigma}_s$ and $\hat{\sigma}_b$, and approximately constant for each observer.

Estimates of the Pitch Processor Noise.

Table 9.2 presents estimates of σ_p and $\sqrt{\hat{\sigma}_p^2 \times (2/u)}$ from the various experiments described above; these estimates are displayed graphically as a function of $\hat{\sigma}_s$ in figures 9.2 and 9.3. Also presented in the table, and in figures 9.4 and 9.5 are estimates of the difference $\sqrt{\hat{\sigma}_p^2 - \hat{\sigma}_s^2}$, which would be expected to remain constant for each observer if stimulus uncertainty had no effect upon performance. This analysis considers only those signals containing the 4th and 5th harmonics of 200 Hz.

Figure 9.2. Estimates of the processor noise σ_p as a function of $\hat{\sigma}_s$. Based on Experiments 5 and 6. Note the linear scale. Filled symbols, observer A.F., empty symbols, observer G.S.

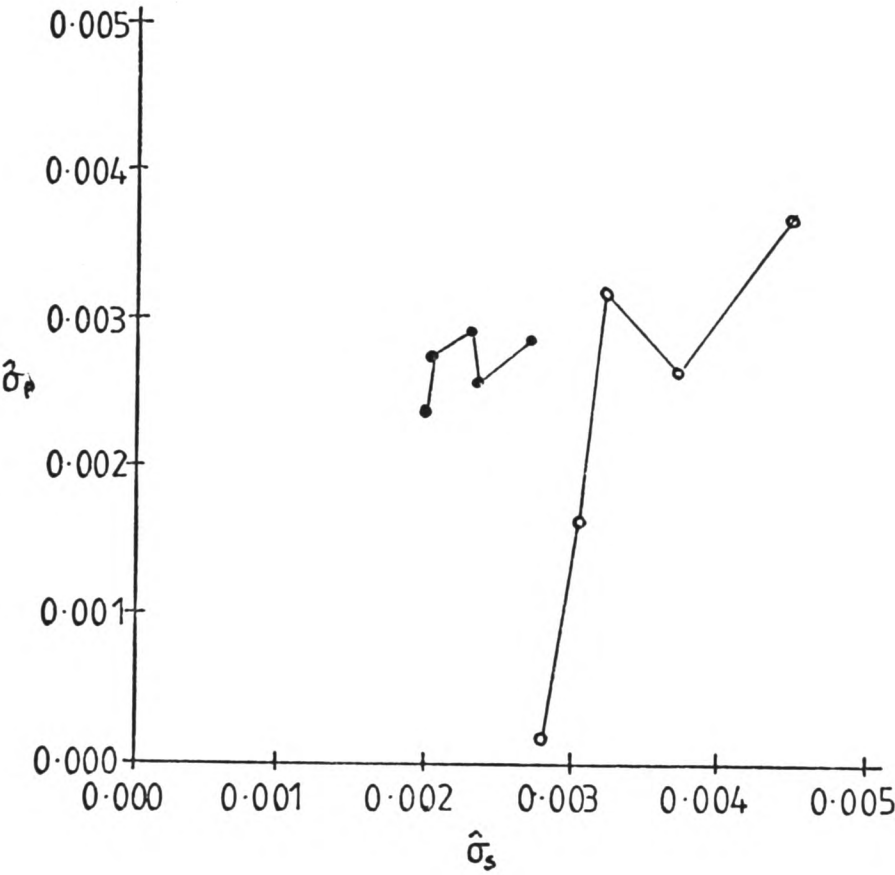


Figure 9.3. Estimates of the processor noise $\hat{\sigma}_p$ as a function of $\hat{\sigma}_s$. Here the scales are logarithmic. Data from Experiments 8 and 10, 9 and 10, and 12 and 13.

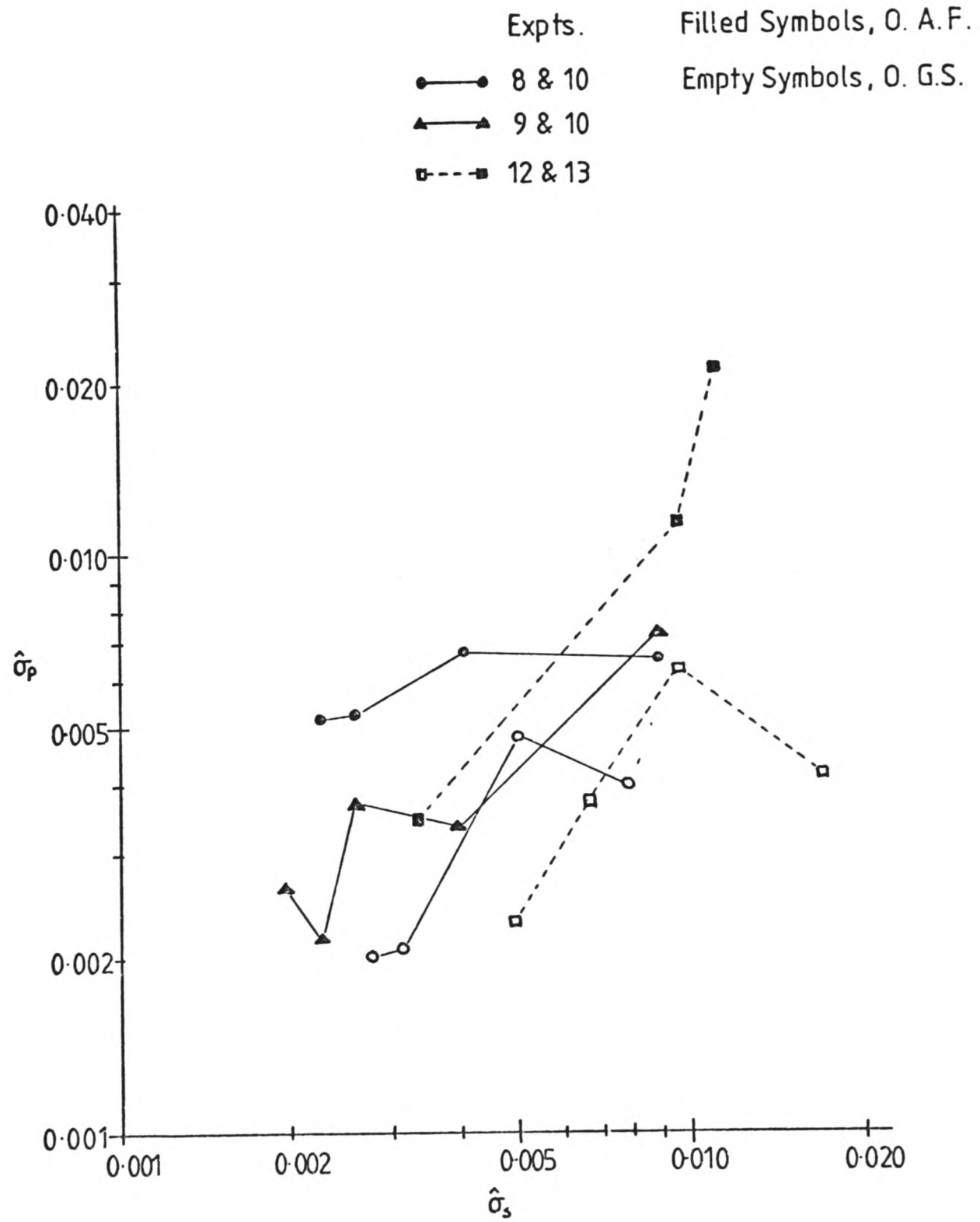


Figure 9.4. Estimates of the difference $\sqrt{(\hat{\sigma}_0^2 - \hat{\sigma}_s^2)}$ as a function of $\hat{\sigma}_s$. Data from Experiments 5 and 6. The scales are linear. Filled symbols, observer A.F., empty symbols, observer G.S.

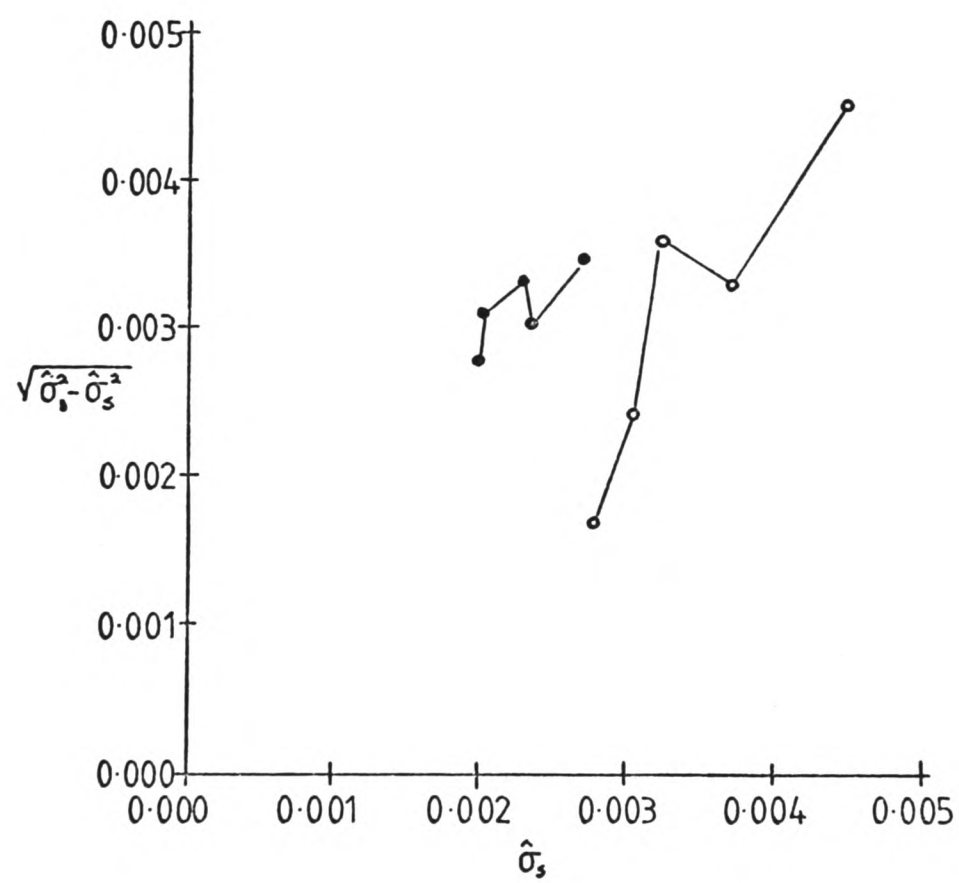


Figure 9.5. Estimates of the difference $\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}$ as a function of $\hat{\sigma}_s^2$. Data from Experiments 8 and 10, 9 and 10, and 12 and 13. The scales are logarithmic.

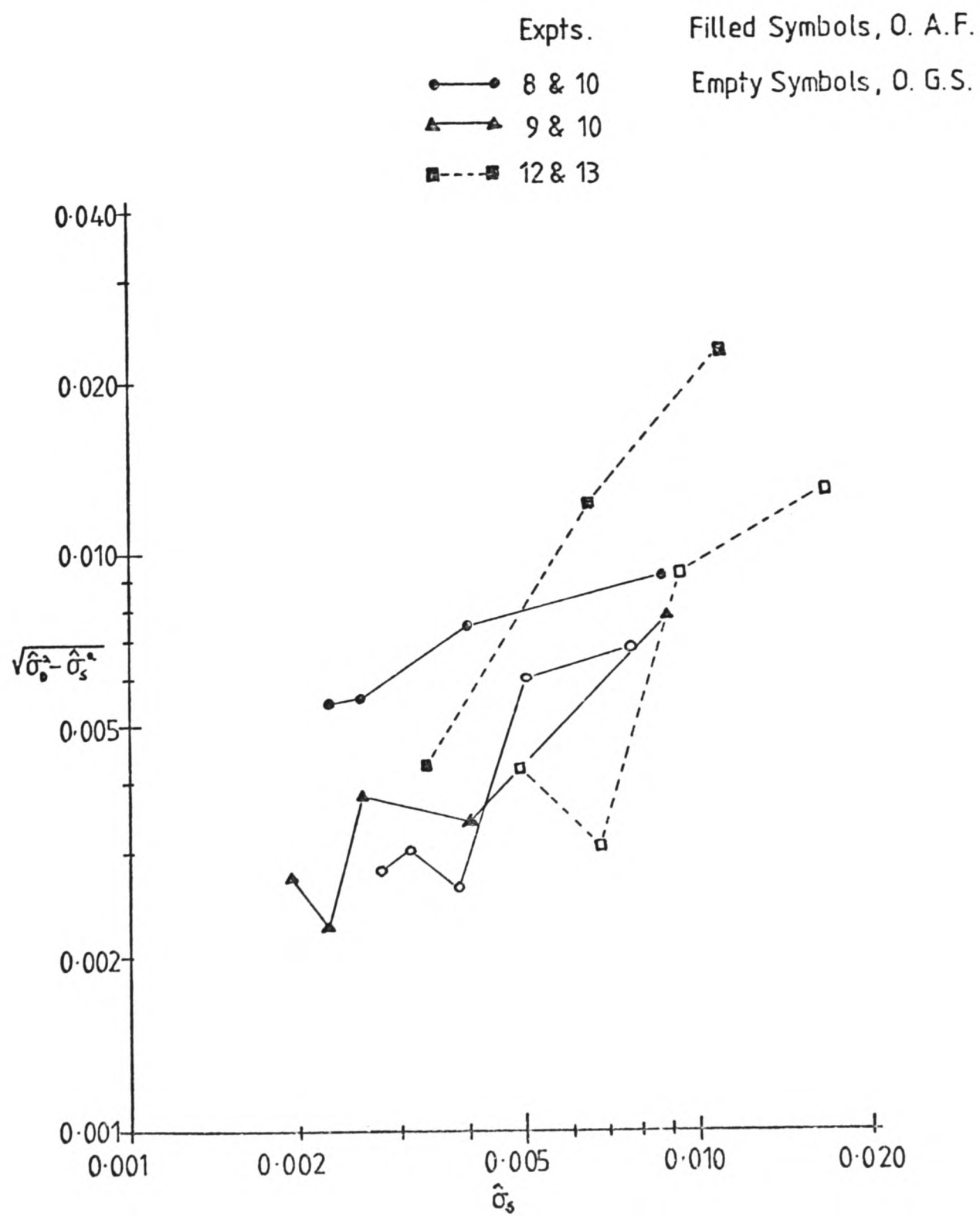


Table 9.2, Estimates of σ_p , $\sqrt{\hat{\sigma}_p^2 \cdot (2/u)}$ and $\sqrt{\hat{\sigma}_D^2 - \hat{\sigma}_S^2}$. $r(x,y)$ is the product-moment correlation between $\underline{\sigma}_S$ and the indicated estimate. The entry for the source of the estimates refers to the experiment from which the estimates are derived.

Obs. Source	SL dB	15	20	30	40	50
A.F. 6 and 5						
	$\hat{\sigma}_s$	0.00272	0.00231	0.00200	0.00203	0.00236
	$\hat{\sigma}_p$	0.00285	0.00290	0.00238	0.00274	0.00256
u=1.5	$\frac{\sqrt{\hat{\sigma}_p^2(2/u)}}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00329	0.00323	0.00275	0.00316	0.00296
	$\frac{r(\hat{\sigma}_s, \hat{\sigma}_p)}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00344	0.00333	0.00277	0.00310	0.00305
	$r(\hat{\sigma}_s, \hat{\sigma}_p) = 0.553$					
	$r(\hat{\sigma}_s, \sqrt{\hat{\sigma}_b^2 - \hat{\sigma}_s^2}) = 0.458$					
G.S. 6 and 5						
	$\hat{\sigma}_s$	0.00372	0.00305	0.00278	0.00323	0.00451
	$\hat{\sigma}_p$	0.00263	0.00163	0.00064	0.00317	0.00367
u=1.5	$\frac{\sqrt{\hat{\sigma}_p^2(2/u)}}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00304	0.00188	0.00074	0.00366	0.00424
	$\frac{r(\hat{\sigma}_s, \hat{\sigma}_p)}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00329	0.00241	0.00167	0.00360	0.00452
	$r(\hat{\sigma}_s, \hat{\sigma}_p) = 0.733$					
	$r(\hat{\sigma}_s, \sqrt{\hat{\sigma}_b^2 - \hat{\sigma}_s^2}) = 0.828$					
Obs. Source	E-N ₀ dB	20	25	30	35	40
A.F. 8 and 10						
	$\hat{\sigma}_s$	0.00888	0.00404	0.00261	0.00228	0.00197
	$\hat{\sigma}_p$	0.00653	0.00688	0.00527	0.00522	-
u=1.5	$\frac{\sqrt{\hat{\sigma}_p^2(2/u)}}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00754	0.00794	0.00609	0.00603	-
	$\frac{r(\hat{\sigma}_s, \hat{\sigma}_p)}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00906	0.00745	0.00558	0.00546	-
	$r(\hat{\sigma}_s, \hat{\sigma}_p) = 0.642$					
	$r(\hat{\sigma}_s, \sqrt{\hat{\sigma}_b^2 - \hat{\sigma}_s^2}) = 0.952$					
G.S. 8 and 10						
	$\hat{\sigma}_s$	0.00783	0.00506	0.00384	0.00315	0.00278
	$\hat{\sigma}_p$	0.00397	0.00485	(1)	0.00206	0.00199
u=1.5	$\frac{\sqrt{\hat{\sigma}_p^2(2/u)}}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00458	0.00537		0.00238	0.00230
	$\frac{r(\hat{\sigma}_s, \hat{\sigma}_p)}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.00681	0.00603	0.00261	0.00303	0.00280
	$r(\hat{\sigma}_s, \hat{\sigma}_p) = 0.719$					
	$r(\hat{\sigma}_s, \sqrt{\hat{\sigma}_b^2 - \hat{\sigma}_s^2}) = 0.904$					
A.F. 9 and 10						
	$\hat{\sigma}_s$	0.00888	0.00404	0.00261	0.00228	0.00197
	$\hat{\sigma}_p$	0.00741	0.00335	0.00367	0.00217	0.00263
u=1	$\frac{\sqrt{\hat{\sigma}_p^2(2/u)}}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$	0.01048	0.00474	0.00519	0.00307	0.00372
	$\frac{r(\hat{\sigma}_s, \hat{\sigma}_p)}{\sqrt{(\hat{\sigma}_b^2 - \hat{\sigma}_s^2)}}$					
	$r(\hat{\sigma}_s, \hat{\sigma}_p) = 0.965$					

(1) The estimated variance here is negative, and has a value of -5.6577×10^{-7} .

Table 9.2 continued.					
Obs. Source	T, msec	20	25	60	80
A.F. 12 and 13	$\hat{\sigma}_s$	0.01115	0.00654	-	0.00334
	$\hat{\sigma}_p$	0.02114	0.01128	-	0.00349
	$u=1.5 \sqrt{\hat{\sigma}_p(2/u)}$	0.02441	0.01303	-	0.00403
	$\sqrt{(\hat{\sigma}_0^2 - \hat{\sigma}_s^2)}$	0.02256	0.01219	-	0.00421
	$r(\hat{\sigma}_s, \hat{\sigma}_p) = 0.999$				
	$r(\hat{\sigma}_s, \sqrt{\hat{\sigma}_0^2 - \hat{\sigma}_s^2}) = 0.999$				
G.S. 12 and 13	$\hat{\sigma}_p$	0.01720	0.00943	0.00680	0.00495
	$\hat{\sigma}_s$	0.00413	0.00628	0.00371	0.00229
	$u=1.5 \sqrt{\hat{\sigma}_p(2/u)}$	0.00477	0.00725	0.00423	0.00264
	$\sqrt{(\hat{\sigma}_0^2 - \hat{\sigma}_s^2)}$	0.01284	0.00916	0.00306	0.00418
	$r(\hat{\sigma}_s, \hat{\sigma}_p) = 0.319$				
	$r(\hat{\sigma}_s, \sqrt{\hat{\sigma}_0^2 - \hat{\sigma}_s^2}) = 0.943$				

The values shown in table 9.2 are of course subject to considerable error of estimation, dependent upon the standard error of the estimates of $\underline{\sigma}_s$ and $\underline{\sigma}_p$, which is approximately proportional to the mean of the estimate.

The Fit of the Modified Model to the Discrimination Data.

From figure 9.5, it is apparent that the difference $\sqrt{(\hat{\sigma}_0^2 - \hat{\sigma}_s^2)}$ is not constant, but is an increasing function of $\underline{\hat{\sigma}_s}$. The slope of this function is approximately 1 for all the comparisons shown in the figure, and the correlation between $\underline{\hat{\sigma}_s}$ and $\sqrt{(\hat{\sigma}_0^2 - \hat{\sigma}_s^2)}$ is typically large. Over all the tabulated estimates, the correlation is 0.7545. The estimates of $\underline{\sigma}_p$ shown in figure 9.2 and 9.3, with the exception of the values derived from observer A.F.'s performance as a function of signal duration, are approximately constant and show little dependence

on $\hat{\sigma}_s$. The overall correlation of $\hat{\sigma}_s$ and $\hat{\sigma}_r$ is 0.4952. The values of the estimates of σ_r lie between about 0.002 and 0.007. A Structural Relation (Kendall and Stuart, 1961), which gives maximum-likelihood estimates of the slope and intercept of a relationship, has been fitted to these estimates, and the results of these analyses are shown in table 9.3.

Table 9.3, Structural Relation analysis of the dependence of $\hat{\sigma}_r$ and $\sqrt{\hat{\sigma}_s^2 - \hat{\sigma}_s^2}$ upon $\hat{\sigma}_s$.						
Relation	Slope	Intercept	r(x,y)	t	df	p
a) $\hat{\sigma}_s$ and $\hat{\sigma}_r$.	0.8939	-2.2845 x 10-6	+0.4952	2.716	27	<0.02
b) $\hat{\sigma}_s$ and $\hat{\sigma}_r$.	0.2541	2.3962 x 10-3	+0.4457	2.054	24	n.s.
c) $\hat{\sigma}_s$ and $\sqrt{\hat{\sigma}_s^2 - \hat{\sigma}_s^2}$	1.2213	-5.9068 x 10-4	+0.7545	6.376	28	<0.01
The statistic t and the associated probability tests the difference of the slope of the relationship from a slope of 0. Relation b) was calculated with the estimates from A.F. in Experiments 12 and 13 omitted. The parameter λ , an estimate of the ratio of the variances of the dependent and independent estimates, was given a value 2.0.						

From this Structural Relation analysis, it is clear that the difference $\sqrt{\hat{\sigma}_s^2 - \hat{\sigma}_s^2}$ is dependent upon $\hat{\sigma}_s$, the slope of this relationship being 1.22, which is almost exactly the value expected from the effect of frequency uncertainty. The estimate of σ_r made with the assumption that such uncertainty operates is also dependent upon $\hat{\sigma}_s$, but this dependence is primarily due to estimates from observer A.F. in Experiments 12 and 13, where the standard errors of the estimates of σ_b at short signal durations are much larger than those for other estimates. Such a large error of estimation is likely to result in an overestimation of σ_b , which would account for the strong dependence of these estimates on $\hat{\sigma}_s$ visible in figure 9.3. It

is also possible that at the shorter durations, the pulsed noise in Experiment 12 contributed additional variability, since Weir, Green, Hafter and Burkhardt (1977) found significant detection threshold decrements for 10 msec signals in gated as opposed in continuous masking noise. The omission of this imprecise data results in an improvement in the fit of the model, where the dependence of $\hat{\sigma}_r$ on $\hat{\sigma}_s$ is small, and the best-fitting slope does not differ significantly from 0.

If this modification of Goldstein's model is an appropriate description of the perceptual process involved in residue pitch discrimination, then a Structural Relation analysis of the interdependence of the estimates of σ_D^2 and σ_S^2 derived from the above studies should show a slope which does not differ from the uncertainty factor u . The results of such an analysis appear in table 9.4.

Table 9.4. The dependence of $\hat{\sigma}_0^2$ on $\hat{\sigma}_s^2$.								
Observer	Slope	Intercept	r(x,y)	t	df	p	λ	
A.F. a)	4.5329	-1.4918×10^{-5}	+0.9383	14.866	10	<0.01	47.43	
A.F. b)	1.9270	1.3810×10^{-5}	+0.9628	2.584	7	<0.05	59.44	
A.F. c)	1.6432	3.9852×10^{-6}	+0.9980	14.195	3	<0.01	6.11	
G.S. a)	1.5566	7.1668×10^{-6}	+0.9952	1.324	12	n.s.	10.17	

a) estimates from experiments 5 and 6, 8 and 10 and 12 and 13
b) as a) but omitting estimates from experiments 12 and 13,
c) estimates from experiments 9 and 10.

The slope compared with the data by the statistic t was 1.5 except for the comparison A.F. c), where $\hat{\sigma}_0^2$ was estimated from Experiment 9, and the uncertainty in both tasks was equivalent; here the theoretical slope is 1. The parameter λ was evaluated as $\text{var}(\hat{\sigma}_0^2)/\text{var}(\hat{\sigma}_s^2)$, where $\text{var}(\hat{\sigma}^2)$ was estimated as the variance of a cumulated distribution of the square of $\hat{\sigma}$ derived from 1000 random samples from a normal distribution, having the mean $\hat{\sigma}$ and a standard deviation equal to the standard error of estimation of $\hat{\sigma}$. $\text{Var}(\hat{\sigma}_0^2)/\text{var}(\hat{\sigma}_s^2)$ was then taken to be the average of $\text{var}(\hat{\sigma}_0^2)/\text{var}(\hat{\sigma}_s^2)$ over all the estimates analysed.

This analysis indicates that the modified model fits the performance of observer G.S. rather closely. The internal variance due to the noisy channels or to pitch processing is estimated by the intercept of the Structural Relation, which for this observer is 7.1668×10^{-6} , whose square root $\hat{\sigma}_p$ is 0.002677, which is within the range of the values estimated by the simple subtraction procedure above. $\sqrt{(\hat{\sigma}_p^2 \cdot (2/u))}$, the estimated internal relative standard deviation for each channel, is 0.00309 for this observer. The best-fitting slope for this observer is 1.557, which is very close to the expected slope of 1.5. The fit to the performance of observer A.F. suggests a multiplicative factor larger than that expected from frequency uncertainty. The best-fitting slope for this observer's performance in the pairs of Experiments 5 and 6 and 8 and 10 is

1.927, where uncertainty would predict a slope of 1.5. The fitted slope differs significantly from a slope of 1.5, $t=2.58$, $df=7$, $p<.05$. The best-fitting slope for this observer's performance in Experiments 9 and 10, where uncertainty predicts a slope of 1.0, is 1.643, and this is also a statistically significant difference, $t=14.20$, $df=3$, $p<.01$. Estimates of σ_p from the intercepts of these two fits are 0.003716 and 0.001996. Both of these estimates are well within the range of the estimates from the subtraction procedure. These two estimates correspond to a standard deviation for each noisy channel of 0.00429 and 0.00282 respectively.

There are a number of possible reasons for the rather small but significant deviations between the predictions of the model and the performance of observer A.F. It may be that the variability of the estimates of σ_p lead to estimates which increasingly overestimate this parameter as its true value increases. It is also possible that this observer was subject to a greater degree of stimulus uncertainty than the minimum imposed by the tasks, in which case, the values of u adopted above would be underestimates of the effective value.

The processor noise σ_p estimated here is not directly comparable with the parameter σ estimated by Goldstein (1973), Goldstein et al (1978) and Houtsma (1979), where that parameter refers to the standard deviation due to a single noisy channel.

These latter estimates must include any noise due to external noise, to the peripheral coding of frequency and the effect of stimulus uncertainty, while the estimates derived here exclude these components. It is thus not surprising that the estimates of σ at similar component frequencies made by these other workers are larger at about 0.010 or 0.011 than the estimates made here, which are between 0.00282 and 0.00429. Since these estimates are larger than the typical error of the estimated standard deviations from which the estimate is derived, the inclusion of the internal noise parameter does allow an improved account of performance in the pitch discrimination tasks described above.

CHAPTER 10

Abstract.

This chapter contains a description of two experiments intended to estimate the effects of both signal-to-noise ratio and signal duration upon the relative areas of the major modes of the internal representation of the residue pitch of a two-component harmonic signal. These experiments, Experiments 14 and 15, were essentially similar to each other, and employed the same signals and stimulus conditions as Experiments 8 and 12, which were described in Chapters 7 and 8. These two independent variables were employed as a means of manipulating the standard deviation $\sigma(f'_j)/f_j$; the relations between these independent variables and estimates of this parameter have been described in Chapters 7, 8, and 9.

The obtained results were compared with the predictions derived from computer simulations of the spectral-pattern recognition models due to Goldstein (1973) and to Wightman (1973b) both in terms of the probability of an acceptance of a given pitch match, and in terms of the sensitivity index d' . Although the results were subject to the effects of practice, and to the observer's familiarity with the stimulus material,

performance was generally consistent with the predictions of both models.

Introduction.

If the residue pitch of a two-component complex signal is derived from independent internal representations of the frequencies of the two components, the spectral-pattern recognition models of the perception of residue pitch described in Chapters 1 and 2 can predict the relationship between the standard deviation of the internal representations of the frequencies of the signal components and the relative areas of the several modes of the internal representation of the residue pitch of the signal. This relationship may be quantified by the computer simulation methods described in Chapter 2, the noisiness of the representation of component frequencies being estimated from the discrimination studies described above. The areas of the modes of the internal representation of the residue pitch of the signal can be estimated from the proportions of trials upon which an observer assigns to the signal the pitches corresponding to the frequencies at which these modes lie.

Previously described studies of the distribution of the residue pitch of complex stimuli have either employed a simple matching task where an observer adjusts the frequency of some comparison signal until he is satisfied that the pitch of this signal matches the residue pitch of a standard signal (e.g. Ritsma and Engel, 1964; Schouten et al, 1962), or tasks such as melody or musical interval identification (Goldstein et al, 1978; Houtsma, 1979; Houtsma and Goldsten, 1972). For the purposes of this study, both of these methods have considerable disadvantages. The difficulty with the simple matching method

is that it requires a series of repeated presentations of the standard and comparison signals, the number of presentations being uncontrolled. The sample distributions for the component frequency representations in this task will have a variance which is inversely dependent upon the number of presentations, and which would not be directly comparable with the sample distributions upon which the pitch discrimination judgements from the above 2AFC experiments are assumed to be based. The melody identification task is complex; an incorrect response may indicate an error in assigning a pitch to one or more of the tones comprising the melody, so that the results of this task are hard to analyse. The musical interval identification task is more tractable, but it is unlikely that all of the theoretically possible residue pitches of a given signal would be related to the other signal of the pair by a true musical interval. This task may thus restrict the observer to selecting a residue pitch which does give rise to a musical interval, and hence prevent a response indicating the choice of one or more of the theoretically possible pitches.

The method adopted here employed a Two-Interval-Same-Different (2IAX) task. One of the signals, the standard, was a two-component harmonic signal, expected to have a residue pitch at its fundamental frequency, and perhaps also at other frequencies. The other signal, the comparison, contained a component at its fundamental frequency and had an unambiguous pitch corresponding to that frequency. The use of the terms 'standard' and 'comparison' here refers to the occurrence of several variable comparison signals for each

standard signal, and does not imply any presentation order for the two types of signal. The observer was asked to indicate whether he judged the residue pitch of the standard signal to be the same as, or different from, the pitch of the comparison signal. The fundamental frequency of the comparison signal was equal to a frequency corresponding to the mean of one of the modes of the theoretical pitch distribution of the standard signal, so that the proportion of trials upon which an observer makes the response 'same', given a particular pair of comparison and standard signals, should estimate the probability of his assigning to the standard signal a residue pitch corresponding to the fundamental frequency of the comparison signal. The observer may show a bias in his choice of 'same' and 'different' responses, which would not be present in the simulated predictions. For this reason, the primary comparison of the predicted and obtained results is in terms of the sensitivity index $\underline{d'}$, which is unaffected by such a bias. The adopted task is comparable to the 2IAX task described by Macmillan, Kaplan, and Creelman (1977), where both the standard and comparison signals are variable between trials. Kaplan, Macmillan, and Creelman (1978) have provided a table of values of $\underline{d'}$ for given hit and false alarm rates in such a task.

In order to discourage the use of response cues based upon the degree of extremity of the various comparison frequencies, three standard signal fundamental frequencies were presented randomly. The highest and lowest of these fundamental frequencies were close to the first theoretical pitch modes above and below the fundamental frequency of the intermediate

standard signal.

The Derivation of Predictions.

The presentation conditions for the experiments described below were comparable to those employed in the discrimination Experiments 8 and 12 described above, where the two signals whose pitches were compared contained components from different frequency regions. It was argued in the previous chapter that in such a task, the effective variance of the internal representations of component frequency includes an internal noise component, and that the external and internal noise contributions were dependent upon the degree of frequency uncertainty. Therefore, the predicted pitch pdfs for the current experiments should be derived from distributions of component frequency estimates which include both the internal noise term and an uncertainty factor. The internal noise component is included in the estimates of $\sigma(f'_o)/f_o$ from Experiments 8 and 12, so that the estimates of $\sigma(f'_j)/f_j$ for the computer simulation of the pitch pdfs may be based upon these estimates of $\sigma(f'_o)/f_o$. The uncertainty in the experiments described below is greater than that in Experiments 8 and 12. In the experiments described below, six two-component complexes are presented in a random order. These complexes are so constructed that their components will lie in two of five possible frequency regions. The component frequencies are given in table 10.1,; five critical band filters centered at about 680, 830, 1025, 1260 and 1540 Hz should suffice for the encoding of these components.

Table 10.1. Component frequencies of the two-component signals presented in Experiments 14 and 15.			
f_o	n	$f_o \cdot n$	$f_o \cdot (n + 1)$
170	4	680	850
170	5	850	1020
205	4	820	1025
205	5	1025	1230
258	4	1032	1290
258	5	1290	1540

The uncertainty for the current experiments thus exceeds that for Experiments 8 and 12 by a factor of 5/3, and the estimates of $\sigma(f'_j)/f_j$ for the computer simulations include this factor. If both components of each signal are represented internally with equal $\sigma(f'_j)/f_j$, then the effective value of $\sigma(f'_j)/f_j$ may be simply estimated by applying the factor $\sqrt{(5/3)} \times 2$ to the estimates of $\sigma(f'_o)/f_o$ for the same signal from Experiments 8 and 12. For observer G.S., the obtained estimates of $\sigma(f'_j)/f_j$ from Experiment 3 are approximately equal for each component, and the observed estimates of $\sigma(f'_o)/f_o$ for $n=4$ and $n=5$ from Experiments 5, 8 and 12 are also approximately equal. The predictions for this observer were thus derived with the effective $\sigma(f'_j)/f_j$ derived as just stated. For observer A.F. however, there is evidence from Experiment 7 that the 4th harmonic was represented with a relative standard deviation about 1.5 times larger than that for the 5th harmonic. For this observer then, the effective $\sigma(f'_j)/f_j$ was estimated as above, but for the signal having $n=4$, the estimates were adjusted in accordance with the observed factor of 1.5. The estimates of $\sigma(f'_j)/f_j$ entered into the simulation procedure are given in table 10.2.

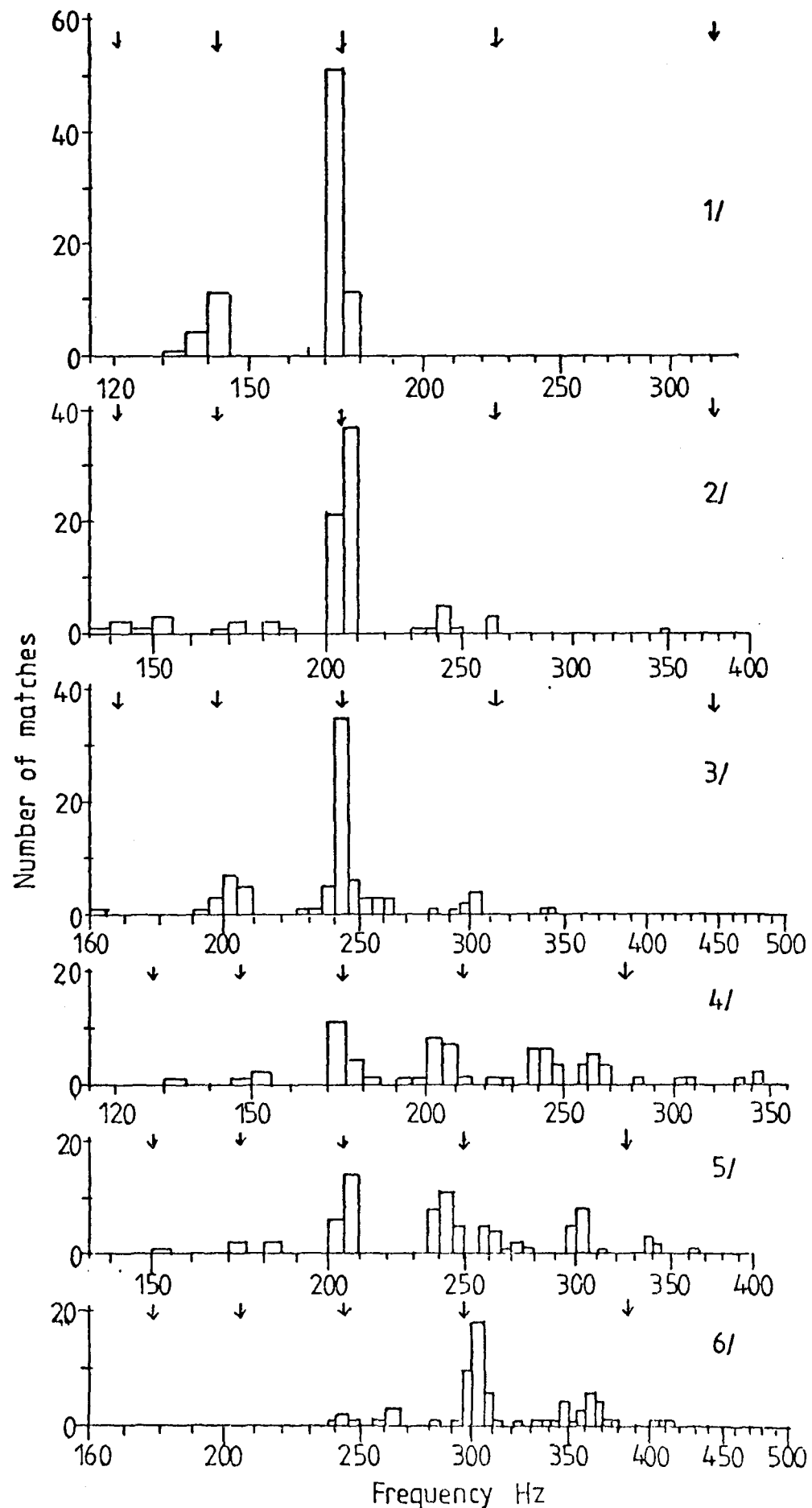
Table 10.2. Estimates of $\sigma(f'_j)/f_j$ for the simulation of predicted residue pitch pdfs. The legend n refers to the lower harmonic number of the standard signal, and the legend m to the harmonic to which $\sigma(f'_j)/f_j$ is appropriate.

Obs	E-N ₀ dB	n	m	$\sigma(f'_j)/f_j$	n	m	$\sigma(f'_j)/f_j$
A.F.	20	4	4	0.020055	5	5 & 6	0.012769
			5	0.013370			
	25	4	4	0.013404	5	5 & 6	0.006703
			5	0.008936			
	30	4	4	0.009755	5	5 & 6	0.006225
			5	0.006503			
	35	4	4	0.009352	5	5 & 6	0.005931
			5	0.006235			
	G.S.	4	4 & 5	0.013398	5	5 & 6	0.019521
			4 & 5	0.010159			
			4 & 5	0.005646			
			4 & 5	0.005087			
Obs	T ₀ msec	n	m	$\sigma(f'_j)/f_j$	n	m	$\sigma(f'_j)/f_j$
A.F.	20	4	4	0.039756	5	5 & 6	0.028725
			5	0.026504			
	25	4	4	0.021876	5	5 & 6	0.015744
			5	0.014584			
	40	4	4	0.014639	5	5 & 6	0.011606
			5	0.009760			
	60	4	4	0.012287	5	5 & 6	0.008874
			5	0.008191			
	80	4	4	0.008496	5	5 & 6	0.005604
			5	0.005664			
	100	4	4	0.009744	5	5 & 6	0.006225
			5	0.006496			
G.S.	20	4	4 & 5	0.027712	5	5 & 6	0.029764
	25	4	4 & 5	0.016971	5	5 & 6	0.015093
	40	4	4 & 5	0.010580	5	5 & 6	0.012354
	60	4	4 & 5	0.009627	5	5 & 6	0.010984
	80	4	4 & 5	0.008362	5	5 & 6	0.009556
	100	4	4 & 5	0.005990	5	5 & 6	0.007690

The Matched Pitches of the Two-Component Signals: A Pilot Study.

A pilot study, where each observer matched the pitch of each of the six standard signals, was performed to check that the theoretical pitch modes conformed to those selected by the observers. Here, the observer heard alternating 100 msec bursts of one of the six standard signals and a sinusoid. The standard signal, but not the sinusoidal comparison signal, was presented with the sum of a white noise and a low-pass filtered noise. The spectral density of the white noise was varied between 20 and 40 dB below the level of the two components of the standard signal. The observer varied the frequency of the sinusoid until he was satisfied with the pitch match, at which point the experimenter noted the frequency of the sinusoid, and started the next trial. Observer A.F. made rather few matches at frequencies other than the fundamental frequency of the complex signal, while the matches made by observer G.S. were relatively evenly distributed about the theoretical values. The results of this matching procedure for observer G.S. are presented in figure 10.1 together with the theoretical pitch modes.

Figure 10.1. Results of the pitch-matching pilot study for Observer G.S. The figure shows the number of matches for a sinusoidal comparison signal, in bins 5 Hz wide.
 Panel 1/ $\underline{n} = 4$, $f = 170$ Hz.
 Panel 2/ $\underline{n} = 4$, $f = 205$ Hz.
 Panel 3/ $\underline{n} = 4$, $f = 248$ Hz.
 Panel 4/ to 6/, as panels 1/ to 3/ but $\underline{n} = 5$.
 The frequencies corresponding to the means of the theoretical pitch modes for each signal are indicated by downward-pointing arrows. The frequency scale is logarithmic, and scales for the different panels are offset so as to align the true fundamental frequencies of the six two-component signals.



Experimental Method.Observers.

The two observers participating in these two experiments were the same two observers as those participating in the previously described experiments.

Stimuli - Experiment 14.

The standard signal, which was presented in the second observation interval, contained either the 4th and the 5th ($n=4$) or the 5th and the 6th ($n=5$) harmonics of one of three different fundamental frequencies. The fundamental frequencies of the standard signals were chosen so that the highest and lowest frequencies were approximately equal to the frequencies corresponding to the nearest theoretical pitch modes above and below the fundamental frequency mode of the intermediate frequency standard signal. The frequencies corresponding to these two modes differ according to the harmonic numbers of the components of the standard signal, and the actual values employed were the arithmetic means of the frequencies appropriate for the two different pairs of harmonic numbers. The intermediate standard fundamental frequency was 205 Hz, and the other two frequencies were 170 and 258 Hz. For each of the six standard signals, there were five possible comparison frequencies, corresponding to the means of the five most prominent theoretical pitch modes of the standard signal. The comparison signal frequencies for each of the three standard

frequencies and each pair of harmonics are shown in table 10.3. The comparison signal, which was presented in the first observation interval, was identical in structure to the reference signal employed in the frequency discrimination experiments described below.

Table 10.3 Fundamental frequencies f in Hz of the comparison signals presented with each of the standard signals having the lower harmonic number n and fundamental frequency f_o for Experiment 14.					
n	f_o	f	n	f_o	f
4	170	170	5	170	170
		139			144
		220			209
		118			125
		315			271
4	205	205	5	205	205
		167			173
		265			251
		142			150
		379			326
4	258	258	5	258	258
		211			218
		334			316
		174			189
		477			410

Each of the two components of the standard signal were presented at about 30 dB above threshold at 1047 Hz in quiet. Each component of the comparison signal was presented at a level 40 dB above the observer's threshold in quiet at 1047 Hz. The steady-state duration of both the comparison and the standard signals was 87.5 msec, and the signals were gated on and off with an exponential envelope and rise and fall times of 12.5 msec. The inter-signal interval was 200 msec, and the inter-trial interval approximately 3 sec. The comparison signal was presented in quiet, and 100 msec prior to the presentation of the standard signal, the sum of a low-pass filtered noise and a white noise was introduced, the noise spectrum being such that

at 613 Hz, the noise spectral-density was 10 dB below the level of the components of the standard signal, while at frequencies of 800 Hz upwards, the noise spectral-density was 20, 25, 30, 35 or 40 dB below the level of the components. The noise in the region of 613 Hz was present in order to mask any auditory distortion products which might be generated by the two components of the standard signal having $n=4$ and a fundamental frequency of 205 Hz. Because of the variations in the fundamental frequency of the standard signal, these signal-to-noise ratios are appropriate only for standard fundamental frequencies of 205 and 258 Hz. The signal-to-noise ratio for the 4th harmonic of 170 Hz was somewhat smaller, with a maximum of about 30 dB. The noise was gated off 100 msec after the offset of the standard signal. The gating envelope was exponential, and the rise and fall time was 25 msec.

The observer was seated in a sound-insulated booth, and all signals were presented to the observer's left ear through headphones.

Procedure.

A test session of about one hour contained eight blocks of 90 trials. The signal-to-noise ratio was constant throughout each session. Each of the 30 possible signal pairs were presented equally often in each block, in a psuedo-random order. Each observer received two test sessions at each of the signal-to-noise ratios (35, 30, 25 and 20 dB for observer A.F., and 40, 35, 25 and 20 dB for observer G.S.). The order of these

sessions was counterbalanced. Each observer also received a practice session at the highest signal-to-noise ratio prior to the test sessions. The observer was instructed to compare the residue pitch of the standard signal with the pitch of the comparison signal, and to attempt to judge whether these two pitches were the same or different. He responded by writing the letter "S" or "D" on a duplicated response sheet. The observer was told to expect pairs having the same pitch on one fifth of the trials. No feedback was given, but each observer was informed of the general level of his performance when he next attended.

Experiment 15.

Stimuli.

This experiment was similar to Experiment 14, except that the signal-to-noise ratio during the presentation of the standard signal was constant at 30 dB, and the duration of the standard signal was varied. The experiment also tested the possibility that the temporal order of the comparison and standard signals might influence the observer's judgements.

The experiment was run with standard signal durations of 100, 80, 60 and 40 msec with both possible orderings of the comparison and standard signal (Experiment 15a), and also with standard signal durations of 80, 40, 25 and 20 msec with the comparison signal preceding the standard signal (Experiment 15b). The quoted signal duration is the half-power duration of

the signal, the steady-state duration of the signal was 5 msec less than this duration. Both standard and comparison signals were gated on and off with an exponential envelope and rise and fall times of 5 msec. The duration of the comparison signal was always 100 msec; the inter-stimulus interval was 200 msec and the inter-trial interval 2850 msec. During the presentation of the standard signal, the sum of a low-pass filtered noise and a white noise was also present. The white noise had a spectral-density 30 dB below the components of the standard signal, while the spectral-density of the low-pass filtered noise up to 613 Hz was 20 dB below the levels of the components, and at frequencies above this, the low-pass noise level decreased at about 72 dB/octave. The noise was gated on 100 msec before the onset of the standard signal, and gated off 100 msec after its offset, so that the comparison signal was always presented in a silent background.

The fundamental frequencies of the standard signals were the same as those employed in Experiment 12, but only three comparison frequencies were presented for each standard signal; these frequencies are shown in table 10.4.

Table 10.4. Fundamental frequencies f in Hz of the comparison signals presented with each of the standard signals having the lower harmonic number n and fundamental frequency f_0 for Experiment 15.					
n	f	f	n	f	f
4	170	170	5	170	170
		139			144
		220			209
4	205	205	5	205	205
		167			173
		265			251
4	258	258	5	258	258
		211			218
		334			316

Procedure.

Trials were presented in blocks of 90, during which each of the 18 possible signal pairs were presented five times. A test session of about 30 minutes duration contained four blocks, each block having a different standard signal duration. Initially the experiment was run with durations of between 100 and 40 msec (Experiment 15a). Here, each observer received eight test sessions, of which four were given with the comparison signal preceding the standard signal, and four with the alternate ordering. For four of the sessions, the standard signal duration decreased over the four blocks, and for the other four sessions, the signal duration increased over blocks. The order of sessions was counterbalanced with respect to the order of the two signals and the order of change in signal duration. The experiment was run a second time with standard signals of between 80 and 20 msec duration (Experiment 15b). In this case, the comparison signal always preceded the standard signal. The experimental design was otherwise as above, except that each observer received eight test sessions, four with the signal

duration increasing over blocks, and four with that duration decreasing. At least two practice sessions were also given.

The instructions were the same as those for Experiment 14 above, except that the observer was told to expect pairs of the same pitch on one third of the trials.

Results.

The results of the above three experiments will be presented and compared to the predictions from the computer simulation of Goldstein's and Wightman's models in terms of the observed and predicted proportion of "same" responses for each Standard-Comparison pair, and also in terms of the sensitivity index d' . For both the observed and simulated data, d' was obtained from the table presented for the 2IAX task by Kaplan et al (1978). The hit rate from the observed data was taken as the proportion of "same" responses for standard-comparison pairs of equal fundamental frequency, and the false-alarm rate as the proportion of "same" responses to pairs of differing fundamental frequency. For the simulated data, the hit rate was the proportion of trials on which the simulation procedure selected a pitch within the mode of the residue pitch pdf having its mean at or close to the true fundamental frequency, and the false alarm rate as the proportion of trials on which pitch values from other modes were selected.

Analysis of Variance of Observed d' .

The predictions were based upon pitch discrimination data for signals having a fundamental frequency close to 205 Hz, and because pitch discrimination may be affected by fundamental frequency, and also because the signal-to-noise ratio was confounded with fundamental frequency, the predictions may be inappropriate for the standard signals with fundamental frequencies of 170 or 258 Hz. In order to establish the importance of the fundamental frequency of the standard signal, analyses of variance were performed upon the observed values of d' from each observer in each experiment. d' was evaluated from the within-session hit and false alarm rates for each of the six standard signals, and the residual error term was derived from the between-session variability of d' . For Experiment 15 a), where the presentation order of the standard and comparison signals was a between-session factor, d' was computed for each presentation order. Summary tables for the analyses of variance appear as tables 10.5, 10.6 and 10.7.

Table 10.5 Analysis of variance of \bar{d}' for Experiment 14. Each \bar{d}' value is based upon 15 trials, and the residual error is based upon two sessions. Probabilities of greater than 0.05 are not shown.

Observer A.F.

Source	Sum of squares	df	Mean square	F	P
f_o	5.633713	2	2.816857	2.0065	
n	0.856003	1	0.856003	0.6098	
$E-N_o$	23.326528	3	7.775509	5.5387	<0.010
$n \times f_o$	1.126504	2	0.563252	0.4012	
$E-N_o \times f_o$	6.790621	6	1.131770	0.8062	
$E-N_o \times n$	1.282572	3	0.427524	0.3045	
$E-N_o \times n \times f_o$	3.968491	6	0.661415	0.4711	
Residual	33.692355	24	1.403848		
Total	76.676788	47			

Observer G.S.

Source	Sum of squares	df	Mean square	F	P
f_o	1.578750	2	0.789375	0.2220	
n	6.120411	1	6.120411	1.7213	
$E-N_o$	7.060518	3	2.353506	0.6619	
$n \times f_o$	25.280616	2	12.640308	3.5550	<0.050
$E-N_o \times f_o$	6.503884	6	1.083981	0.3049	
$E-N_o \times n$	0.713840	3	0.237947	0.0669	
$E-N_o \times n \times f_o$	6.105278	6	1.017546	0.2862	
Residual	85.334806	24	3.555617		
Total	138.698101	47			

Table 10.6. Analysis of variance of $\underline{d'}$ for Experiment 15a). Each $\underline{d'}$ value is based upon 15 trials, and the residual error upon four sessions.

Observer A.F.

Source	Sum of squares	df	Mean square	F	P
f_o	7.230818	2	3.615409	1.8518	
n	0.384313	1	0.384313	0.1968	
order	0.183151	1	0.183151	0.0938	
T_s	27.607066	3	9.202355	4.7135	<0.005
n x f_o	0.667989	2	0.333994	0.1711	
order x f_o	1.752750	2	0.876375	0.4489	
order x n	0.109729	1	0.109729	0.0562	
T_s x f_o	6.057400	6	1.009567	0.5171	
T_s x n	6.173927	3	2.057976	1.0541	
T_s x order	13.337748	3	4.445916	2.2772	
order x n x f_o	0.803760	2	0.401880	0.2058	
T_s x n x f_o	14.532444	6	2.422074	1.2406	
T_s x order x f_o	19.620952	6	3.270159	1.6750	
T_s x order x n	4.802299	3	1.600766	0.8199	
T_s x order x n x f_o	13.206779	6	2.201130	1.1274	
Residual	281.136274	144	1.952335		
Total	397.607421	191			

Observer G.S.

Source	Sum of squares	df	Mean square	F	P
f_o	13.303470	2	6.651735	1.3987	
n	0.486019	1	0.486019	0.1022	
order	4.013636	1	4.013636	0.8440	
T_s	23.792765	3	7.930922	1.6677	
n x f_o	1.356284	2	0.678142	0.1426	
order x f_o	57.040290	2	28.520145	5.9970	<0.005
order x n	10.528136	1	10.528136	2.2138	
T_s x f_o	77.555098	6	12.925850	2.7180	<0.020
T_s x n	120.695328	3	40.231771	8.4597	<0.001
T_s x order	24.749672	3	8.249891	1.7347	
order x n x f_o	30.002219	2	15.001110	3.1543	<0.050
T_s x n x f_o	106.823635	6	17.803940	3.7437	<0.002
T_s x order x f_o	88.883438	6	14.813907	3.1150	<0.010
T_s x order x n	73.130407	3	24.376800	5.1258	<0.005
T_s x order x n x f_o	83.868227	6	13.978037	2.9392	<0.010
Residual	684.820175	144	4.755696		
Total	1401.048898	191			

Table 10.7. Analysis of variance of \bar{d}' for Experiment 15b). Each \bar{d}' value is based upon 15 trials, and the residual error upon eight sessions.

Observer A.F.					
Source	Sum of squares	df	Mean square	F	P
f_o	116.542303	2	58.271150	19.9063	<0.001
n	15.458706	1	15.458706	5.2809	<0.050
T_s	88.942146	3	29.647383	10.1280	<0.001
n x f_o	85.196046	2	42.598023	14.5521	<0.001
T_s x f_o	26.659057	6	4.443177	1.5179	
T_s x n	13.473869	3	4.491290	1.5343	
T_s x n x f_s	35.053284	6	5.842215	1.9958	
Residual	491.781854	168	2.927273		
Total	873.107337	191			
Observer G.S.					
Source	Sum of squares	df	Mean square	F	P
f_o	82.427654	2	41.213827	13.6793	<0.001
n	79.002013	1	79.002013	26.2216	<0.001
T_s	140.960073	3	46.986689	15.5954	<0.001
n x f_o	32.718214	2	16.359107	5.4298	<0.010
T_s x f_o	41.923980	6	6.987330	2.3192	<0.050
T_s x n	11.480561	3	3.826854	1.2702	
T_s x n x f_o	20.282928	6	3.380488	1.1220	
Residual	506.159734	168	3.012856		
Total	914.955234	191			

The predicted and observed values of $\underline{d'}$ for the two experiments are given in full in Appendix A. The effects of theoretical interest in the analyses of variance summarized above are those of signal-to noise ratio and harmonic number in Experiment 14, and signal duration and harmonic number in Experiments 15 a) and 15 b). The absence of other significant main effects, and of interactions of the above factors with others, implies that hit and false alarm rates may be legitimately pooled over other factors to obtain $\underline{d'}$ values representing simply the effects of the two factors of interest, while the presence of any such effects indicates that such pooling is inappropriate. The important outcome of these experiments is the degree of agreement between the observed and predicted results, which is considered below, and the analyses of variance will not be interpreted in detail.

Experiment 14.

The analysis of variance for observer A.F. shows only one significant effect, that of signal-to-noise ratio, $\underline{F}=5.52$ (3,24), $\underline{p}<0.010$. The predicted effect of \underline{n} for this observer is small, and the absence of such an effect upon performance is consistent with the predictions. The performance of observer G.S. shows no significant effect of signal-to-noise ratio, $\underline{F}=0.66$ (3,24) or harmonic number $\underline{F}=1.72$ (1,24), but there is an interaction of standard fundamental frequency and harmonic number, $\underline{F}=3.56$ (2,24), $\underline{p}<0.050$. An Simple Effects analysis (Keppel, 1973) of the effect of harmonic number at each fundamental frequency showed that the effect of harmonic number

was significant for a fundamental frequency of 170 Hz, $F=7.93$, (1,24), $p<0.025$, where the mean d' for $n=4$ is greater than the mean d' for $n=5$ by 0.86. There was no effect of n at the other two fundamental frequencies. Because of this significant interaction, the predictions will be compared with the performance of observer G.S. in this experiment for a standard fundamental frequency of 205 Hz only.

Experiment 15 a)

The performance of observer A.F. shows the predicted effect of signal duration, $F=4.71$ (3,144), $p<0.010$, and no effect of harmonic number, $F=0.197$ (1,144). There were no other significant factors or interactions. Observer G.S. does not show the predicted main effect of duration, $F=1.167$ (3,144), or a main effect of harmonic number, $F=0.102$ (1,144), but both of these factors interact in a complex way with the fundamental frequency of the standard and the presentation order of the standard and comparison signals. The data compared with the predictions for this observer will thus be those for a standard fundamental frequency of 205 Hz only, and performance at each presentation order will be considered separately.

Experiment 15 b).

Observer A.F. shows a significant main effect of signal duration, $F=10.13$ (3,168), $p<0.010$, an effect of harmonic number, $F=5.281$ (1,168) $p<0.025$, and also an interaction of fundamental frequency and harmonic number, $F=19.9$ (2,168)

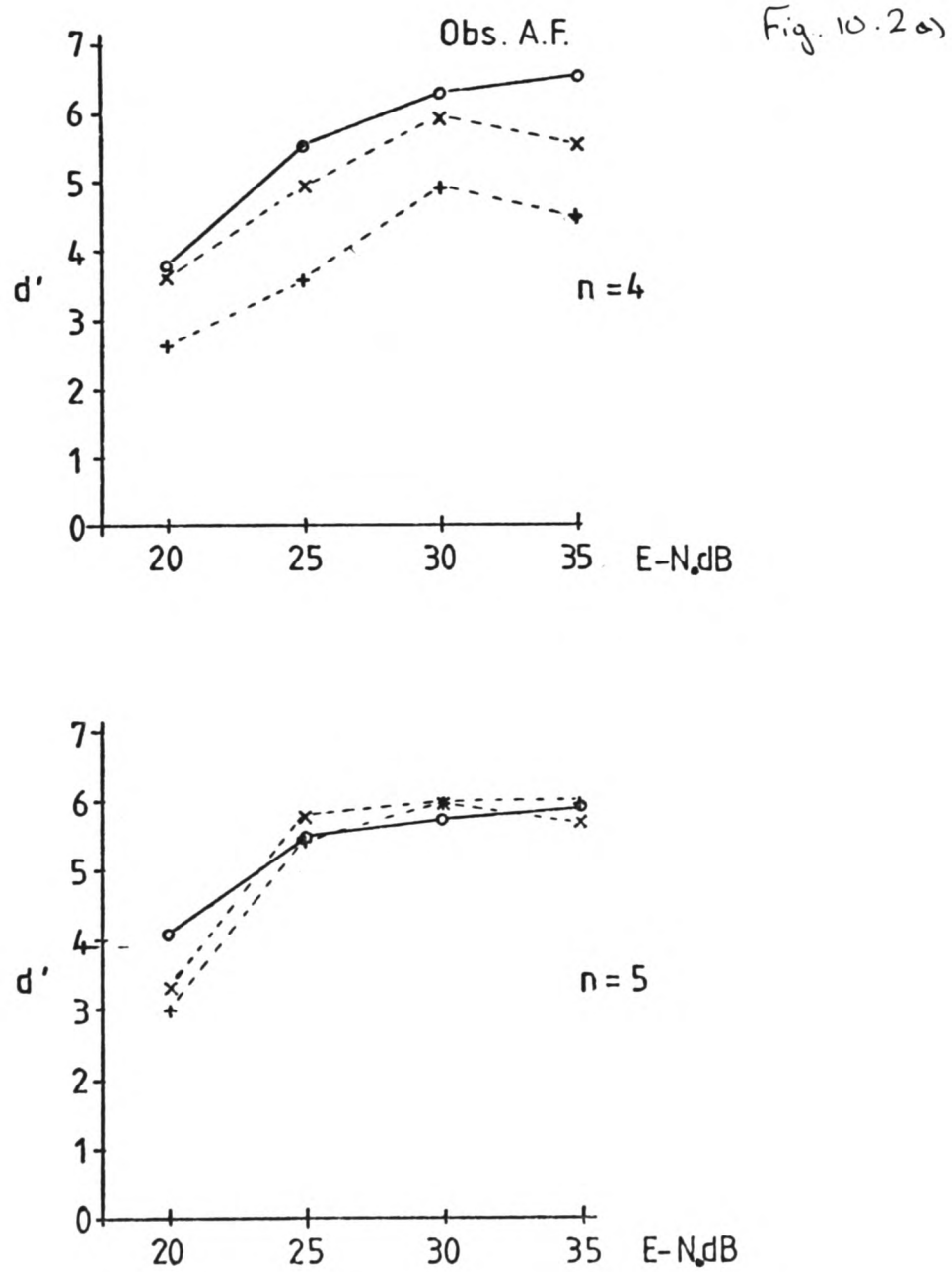
$p < 0.010$. Observer G.S. shows main effects of duration, $F = 15.60$ (3,168) $p < 0.010$, of harmonic number, $F = 26.22$ (1,168), $p < 0.010$, and of fundamental frequency, $F = 13.68$ (2,168) $p < 0.010$. This observer also shows significant interactions of harmonic number with fundamental frequency, $F = 13.68$ (2,168), $p < 0.010$, and of duration with fundamental frequency, $F = 2.319$ (6,168), $p < 0.050$. Since fundamental frequency affects the performance of both observers under the conditions of this experiment, the predictions will be compared with the observers' performance for a standard fundamental frequency of 205 Hz only.

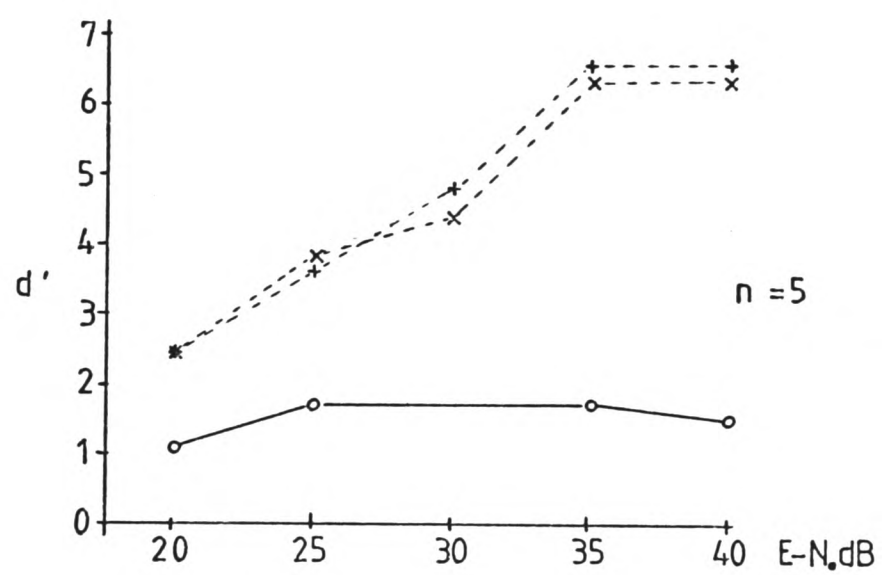
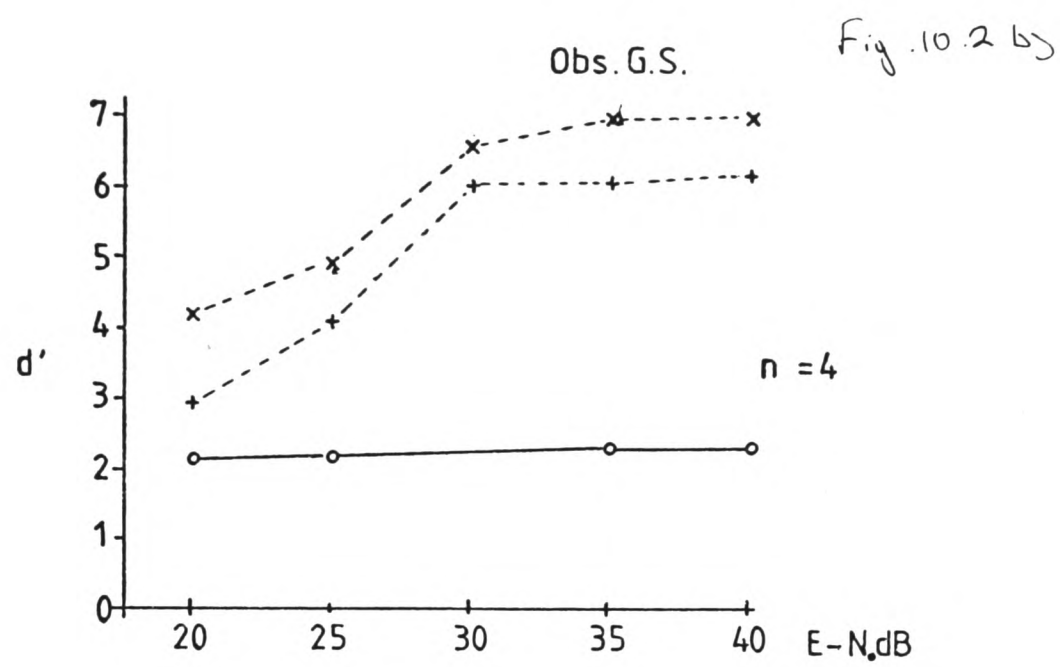
Comparison of Observed and Predicted d' .

Experiment 14.

The observed and predicted values of d' for Experiment 14 are shown in Appendix A and figure 10.2. For observer A.F., the observed d' is derived from hit and false alarm rates averaged over the three standard signal fundamental frequencies. There is rather close agreement between the performance of this observer and the predictions from Goldstein's model; the mean squared deviation of the observed and predicted values of d' ($MSD\ d'$) over the eight conditions is 0.545. The agreement with the predictions from Wightman's model is not so close; here, $MSD\ d'$ is 1.103. For observer G.S., observed values of d' are shown in figure 10.2 for a standard signal fundamental frequency of 205 Hz, and also for proportions pooled over fundamental frequencies. The observed d' for this observer does not show the predicted dependence on signal-to-noise ratio, and the

Figure 10.2. Comparison of observed and simulated d' for Experiment 14; identification of the residue pitch of a two-component harmonic signal as a function of signal-to-noise ratio.
 Panel a), observer A.F.
 Panel b), observer G.S.
 Open symbols and solid lines, observed data for hit and false alarm rates pooled over all three standard signal fundamental frequencies.
 Dashed lines, simulation results -
 x - - - x Goldstein's model
 + - - - + Wightman's model.



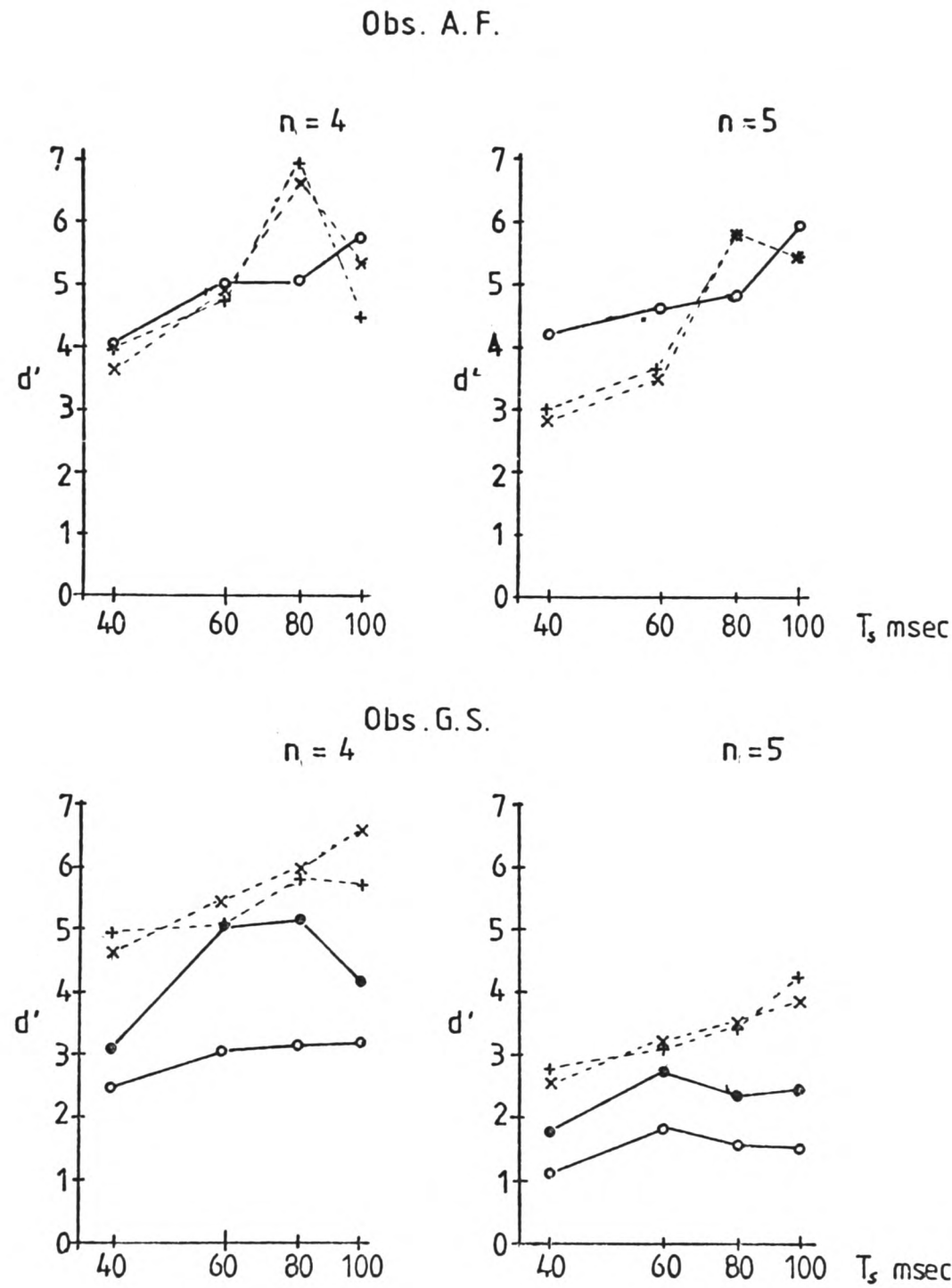


fit of the predictions to the observed data is poor. With a standard signal fundamental frequency of 205 Hz, MSD d' for the predictions from Goldstein's model is 13.883, and 11.418 for Wightman's model.

Experiment 15 a).

The observed and predicted values of d' are given in Appendix A and figure 10.3. For observer A.F., the observed d' is for hit and false alarm rates pooled over all fundamental frequencies and both presentation orders. The predicted d' for this observer at a duration of 100 msec is less than that at 80 msec. The pitch discrimination data on which the prediction for 100 msec is based is taken from a different experiment (Experiment 8) from the discrimination data at the shorter durations, and this feature of the predictions is of no significance. Setting this aside, observer A.F.'s performance is in fairly good agreement with the predictions from both models. For Goldstein's model, MSD d' is 0.870, and for Wightman's model, MSD d' is 1.103. The performance of observer G.S. showed interactions of signal duration and harmonic number with both the fundamental frequency of the standard signal and the presentation order of the standard and comparison signals. The predicted values of d' for this observer are thus compared with observed d' values for a standard frequency of 205 Hz only, and d' is computed separately for each presentation order. The performance of observer G.S. is considerably closer to the predictions than was the case in Experiment 14. For the standard presented before the reference signal, MSD d' is 2.38

Figure 10.3. Comparison of observed and simulated d' for Experiment 15a). Identification of the residue pitch of a two-component harmonic signal as a function of signal duration. Results for two observers. Solid lines and open symbols are for observed d' for data pooled over all three standard fundamental frequencies. Solid lines and filled symbols are for observed d' for a standard fundamental frequency of 205 Hz only. x - - - x simulation of Goldstein's model. + - - - + simulation of Wightman's model.

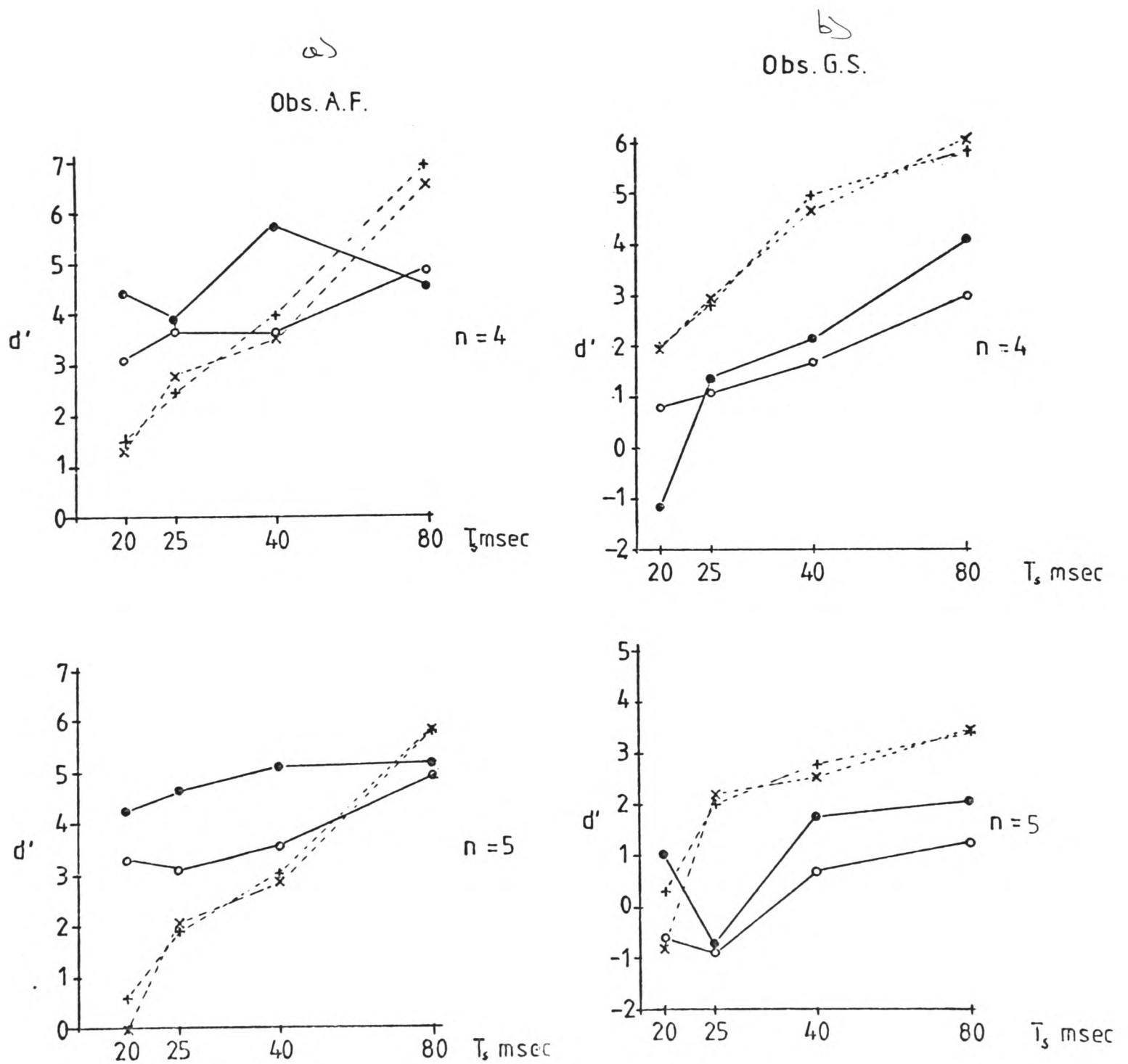


for the predictions from Goldstein's model, and 2.179 for those from Wightman's model. For the reverse presentation order, MSD d' is 1.186 and 0.974 for the two models.

Experiment 15 b).

The observed and predicted values of d' are shown in figure 10.4 and Appendix A. Both observers showed a main effect of standard fundamental frequency in this experiment, and the d' values compared with the predictions are those for a standard fundamental frequency of 205 Hz. The performance of observer A.F. tends to exceed the predicted sensitivity, particularly for signal durations of 20 and 25 msec. MSD d' for this observer is 6.205 for Goldstein's model, and 5.607 for Wightman's. Estimates of $\sigma(f'_j)/f_j$ for the simulated predictions for this experiment were derived from pitch discrimination data from Experiment 12, and it has already been suggested (see Chapter 9) that at short signal durations, the estimate of $\sigma(f'_j)/f_j$ for this observer was an overestimate. If this is indeed the case, the predictions for this experiment will be underestimates of d' , which would account for the observed discrepancy. The observed values of d' for observer G.S. are typically smaller than the predicted values, but show the predicted dependence upon signal duration. Here, MSD d' is 4.511 for Goldstein's model, and 4.178 for Wightman's.

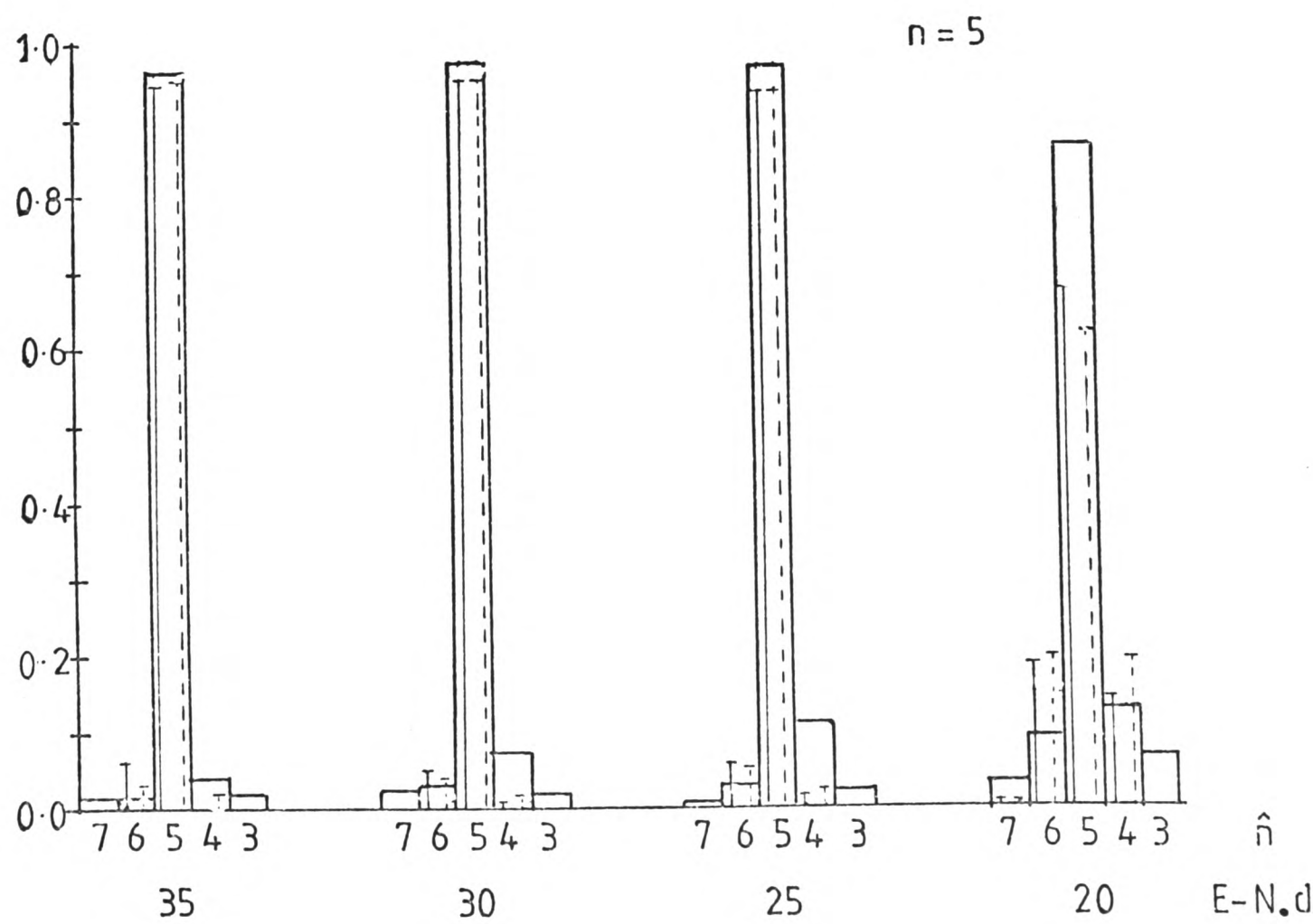
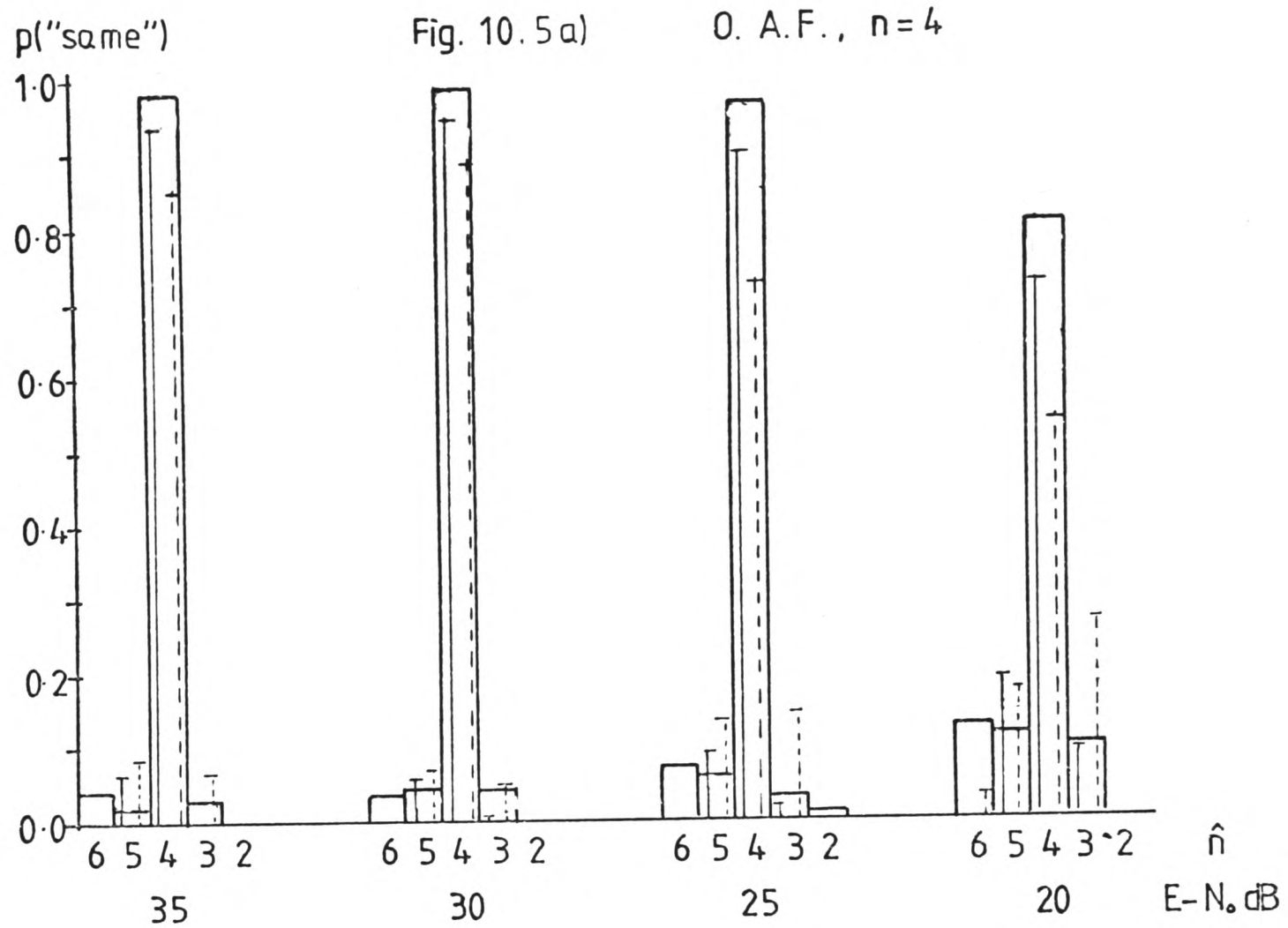
Figure 10.4. Comparison of observed and simulated d' for Experiment 15b); identification of the residue pitch of a two-component harmonic signal as a function of signal duration.
 Panel a), observer A.F., panel b), observer G.S.
 Solid symbols, solid lines - data for standard fundamental frequency of 205 Hz only.
 Empty symbols, solid lines - data for all three standard fundamental frequencies.
 Dashed lines, simulated d' (x - Goldstein's model, + - Wightman's model).



The Observed and Predicted Distribution of "Same"
Responses.

The distribution of "same" responses for the various signal pairs are shown in figures 10.5, 10.6 and 10.7 for Experiments 14, 15a) and 15b) respectively, where the proportions shown are averaged over all three standard fundamental frequencies, and in the case of Experiment 15a), over the two presentation orders. The distributions for a standard fundamental frequency of 205 Hz only appear in figures 10.8, 10.9 and 10.10. Each figure also shows the predicted proportions derived from computer simulation of the models of Goldstein and Wightman. An analysis of this data adds little to the comparison of the observed and predicted values of d' given above, but two points of interest are apparent. The predictions of Goldstein's and Wightman's models are generally very similar except where $\delta(f'_j)/f_j$ exceeds about 0.02, which occurs here for signal durations of 25 and 20 msec in Experiment 15b). Here, Goldstein's model predicts that the probability of a pitch below that corresponding to the fundamental frequency exceeds the probability of a pitch higher than the fundamental frequency, while Wightman's model makes the reverse prediction. The observed proportions for the two observers show different effects, for the judgements of observer A.F. follow the distributions predicted by Goldstein's model, while those of G.S. are comparable to the predictions of Wightman's model. This result does not allow the rejection of one or other model, and Goldstein et al (1978) have shown that the assumption of small constant errors in component frequency estimation allows Goldstein's model to account for

Figure 10.5. Comparison of observed and simulated distributions of "same" responses for Experiment 14; identification of the residue pitch of a two-component harmonic signal as a function of signal-to-noise ratio. The legend \underline{n} indicates the lower harmonic number of the standard signal, and the legend $\hat{\underline{n}}$ the pitch corresponding to the estimated harmonic number $\hat{\underline{n}}$ as in Goldstein's model. Panel a), observer A.F. proportions of 144 trials. Panel b), observer G.S. proportions of 144 trials. The simulation results are shown as lines, the observed data as blocks. The simulation results from Goldstein's model appear as solid lines, and those from Wightman's model as dashed lines.



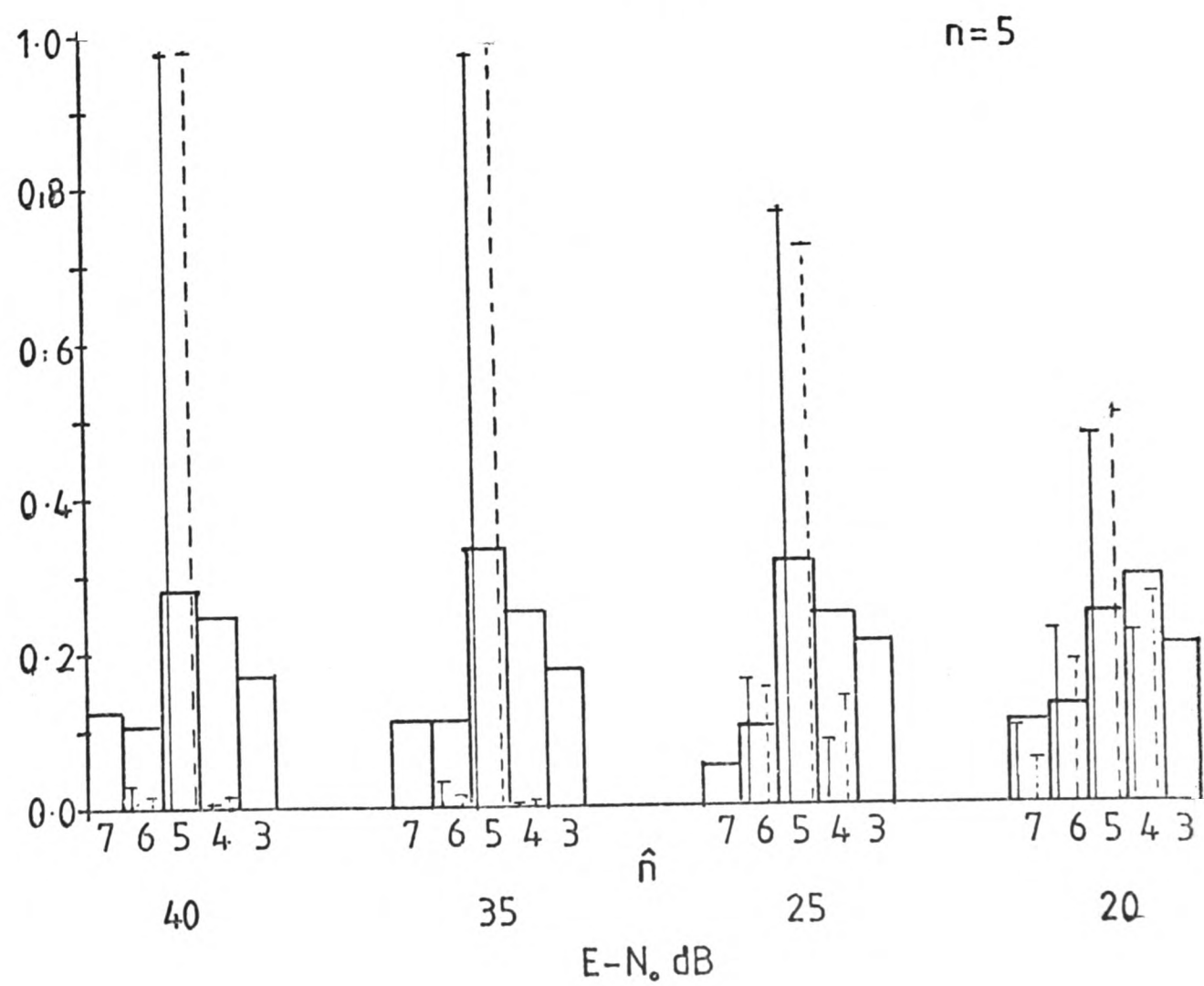
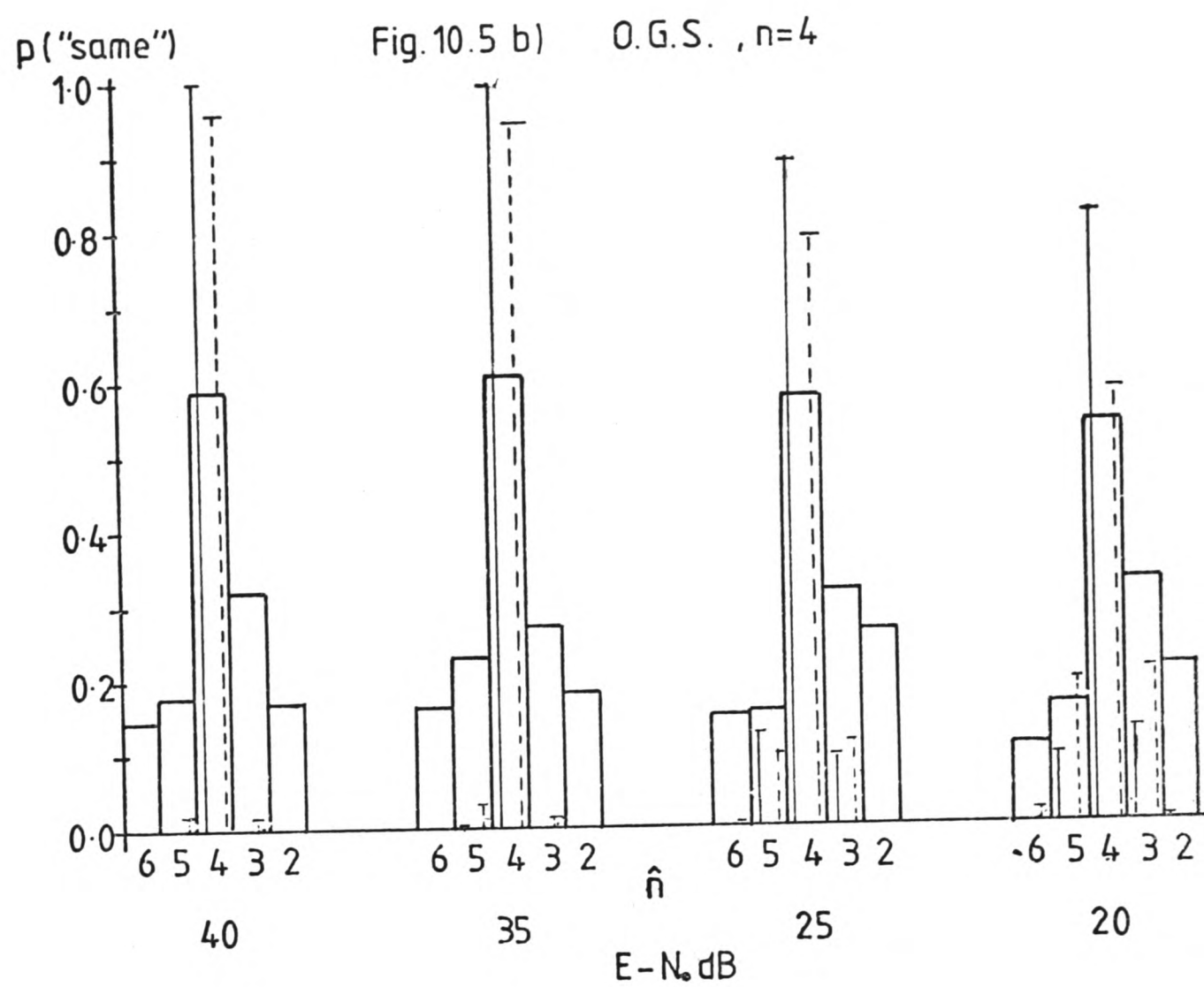


Figure 10.6. Comparison of observed and simulated proportions of "same" responses for Experiment 15a); identification of the residue pitch of a two-component harmonic signal as a function of signal duration. Legend as figure 10.5. The observed data are proportions of 120 trials. Panel a), observer A.F., panel b), observer G.S.

Fig. 10.6 a)

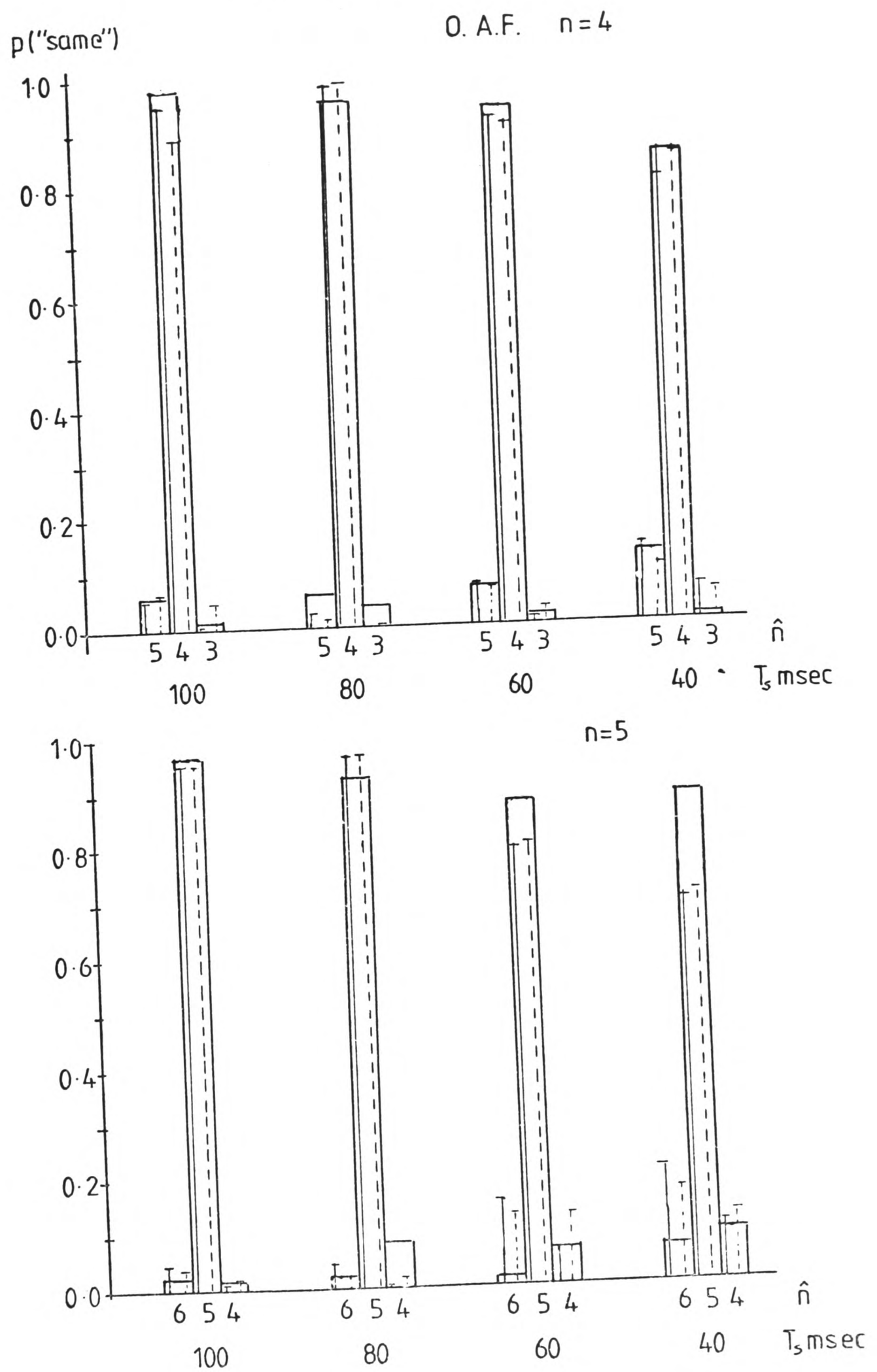


Fig. 10.6 b)

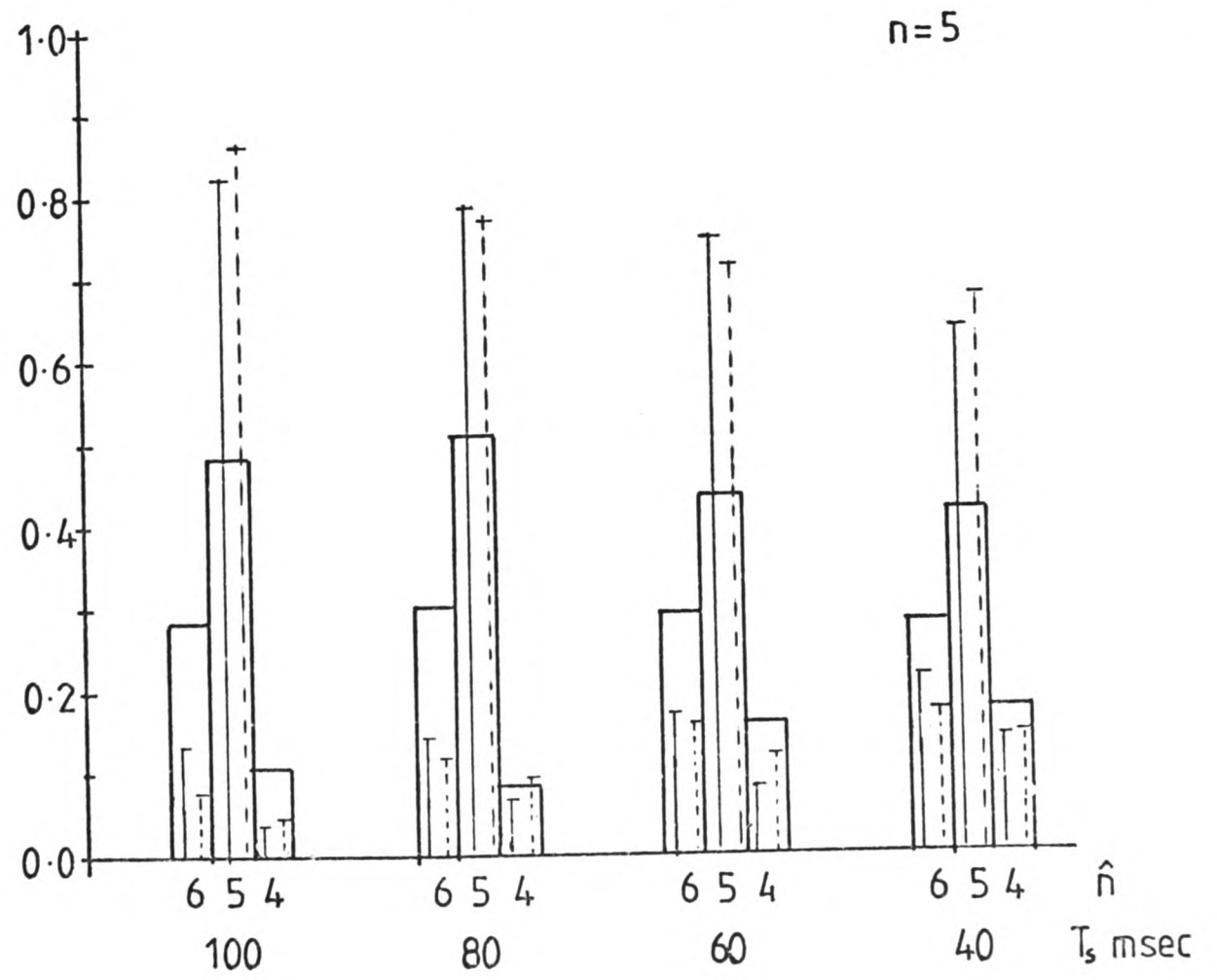
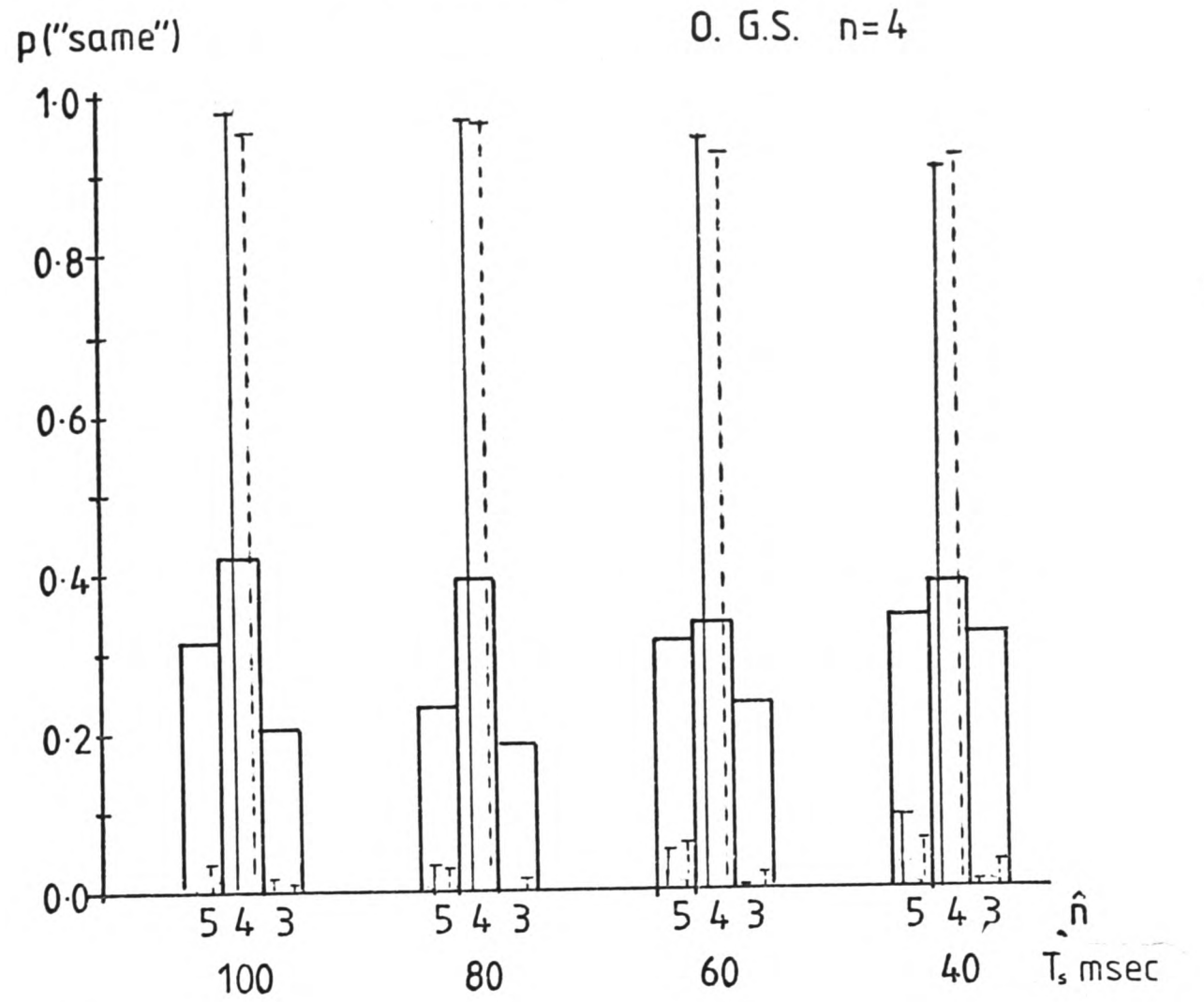
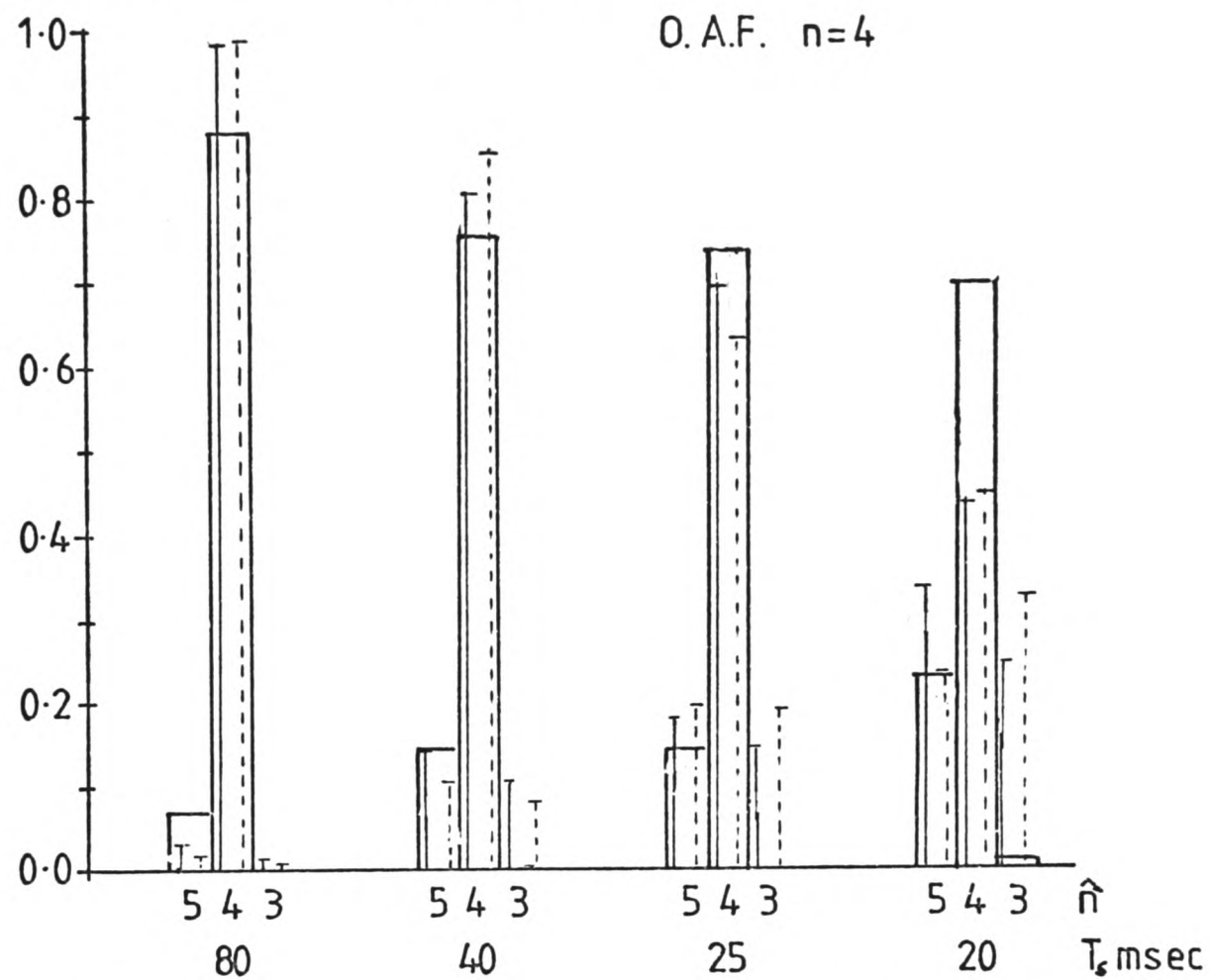


Figure 10.7. Comparison of observed and simulated proportions of "same" responses for Experiment 15b); identification of the residue pitch of a two-component harmonic signal as a function of signal duration. The observed data are proportions of 120 trials. Panel a), observer A.F., panel b); observer G.S. The legends are as for figure 10.5.

$p(\text{"same"})$

Fig. 10.7 a)

O.A.F. $n=4$



$n=5$

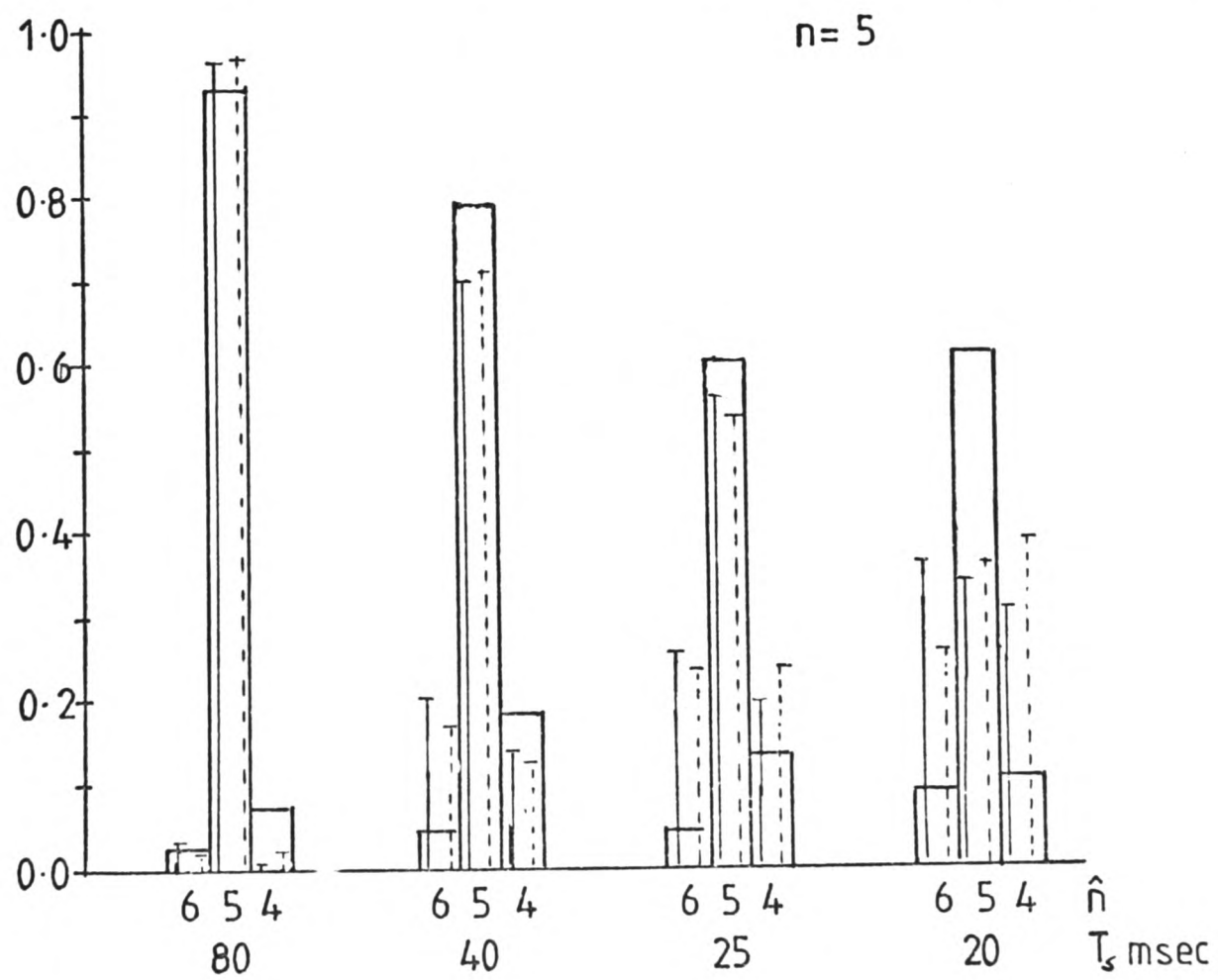
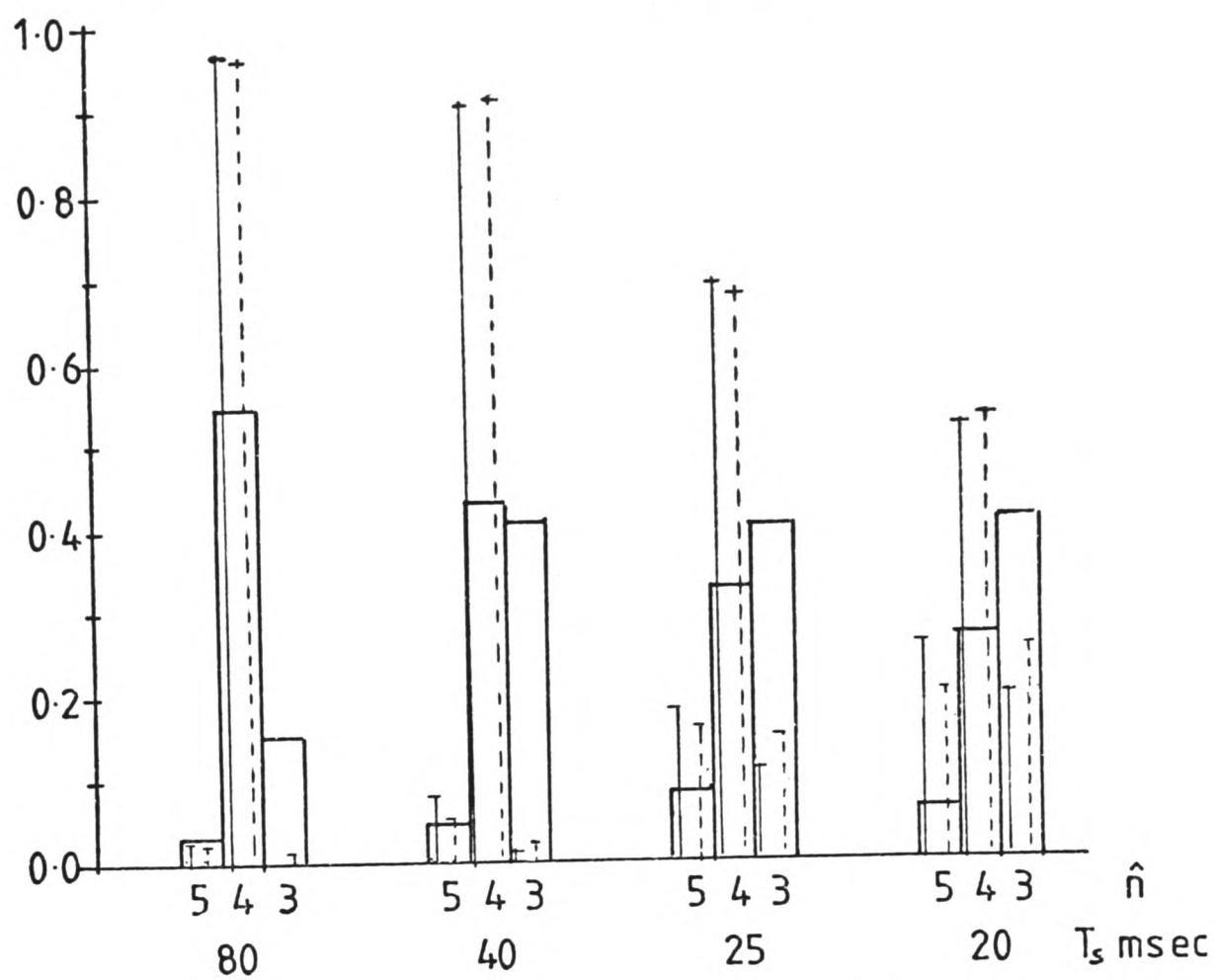


Fig. 10. 7 b)

$p(\text{"same"})$

O. G.S. $n=4$



$n=5$

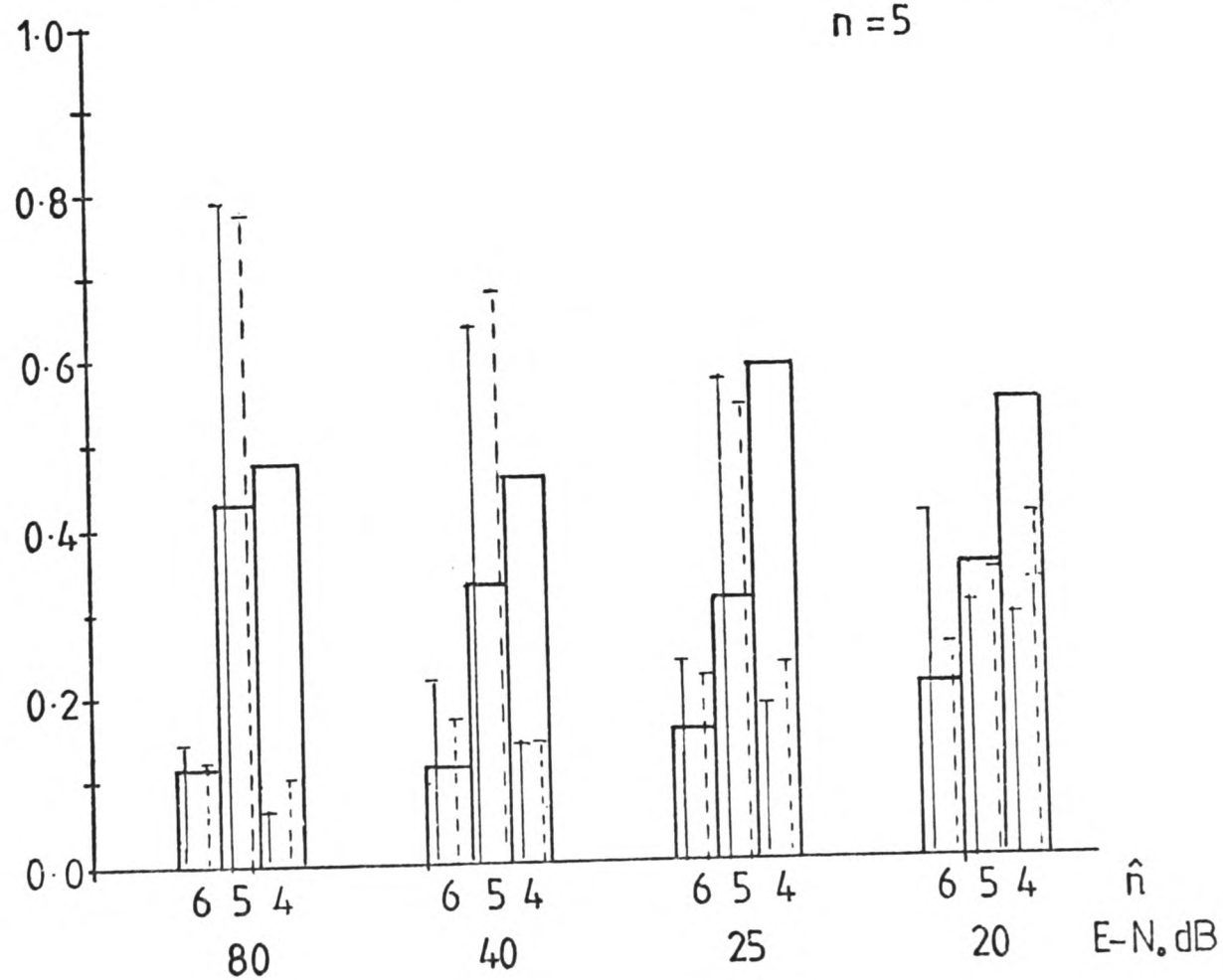
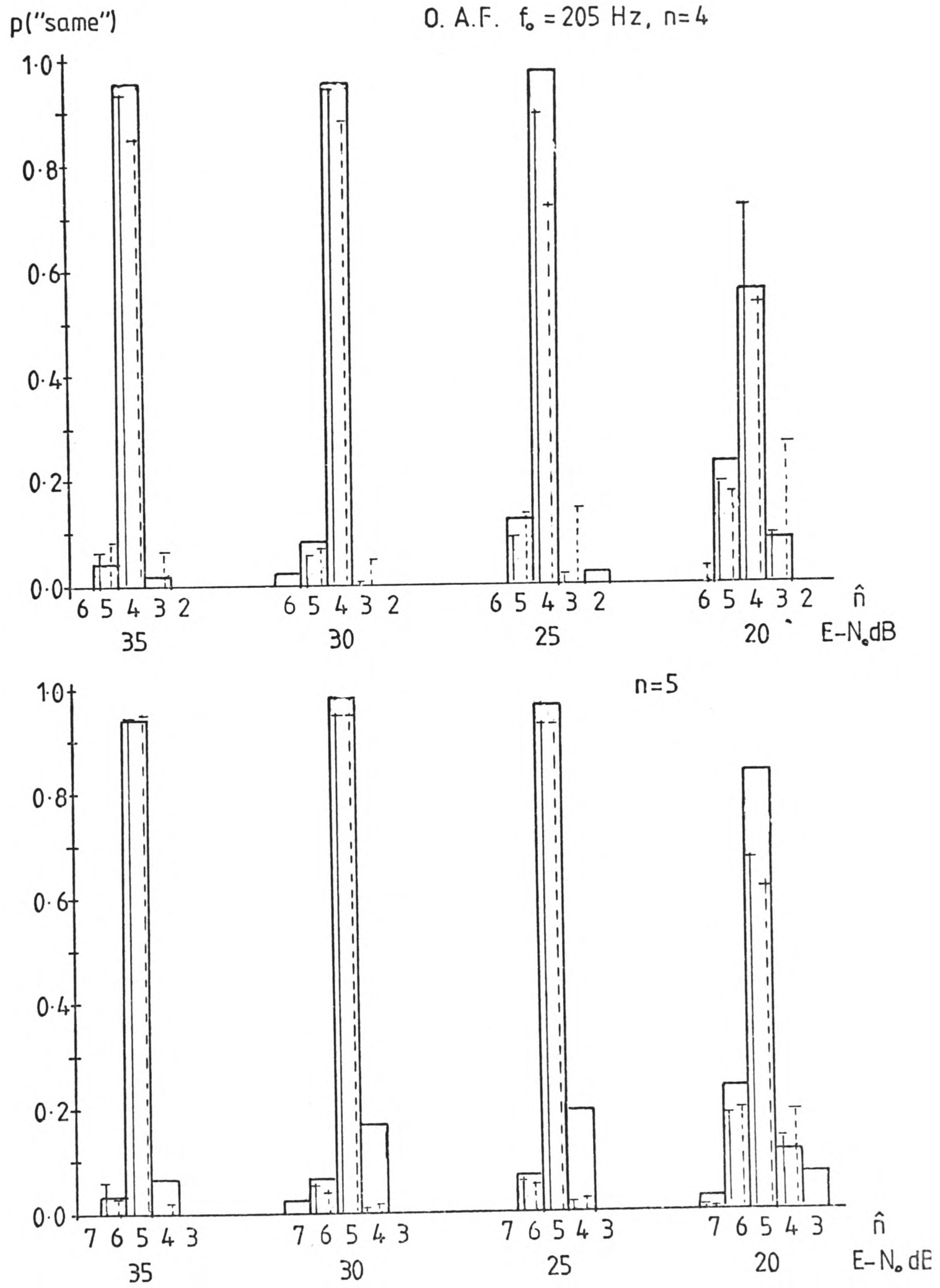


Figure 10.8. Observed distribution of "same" responses for Experiment 14, with a standard fundamental frequency of 205 Hz only. The simulation results are also shown. Panel a), observer A.F., panel b), observer G.S. The observed data are proportions of 48 trials. The simulation results are also shown.

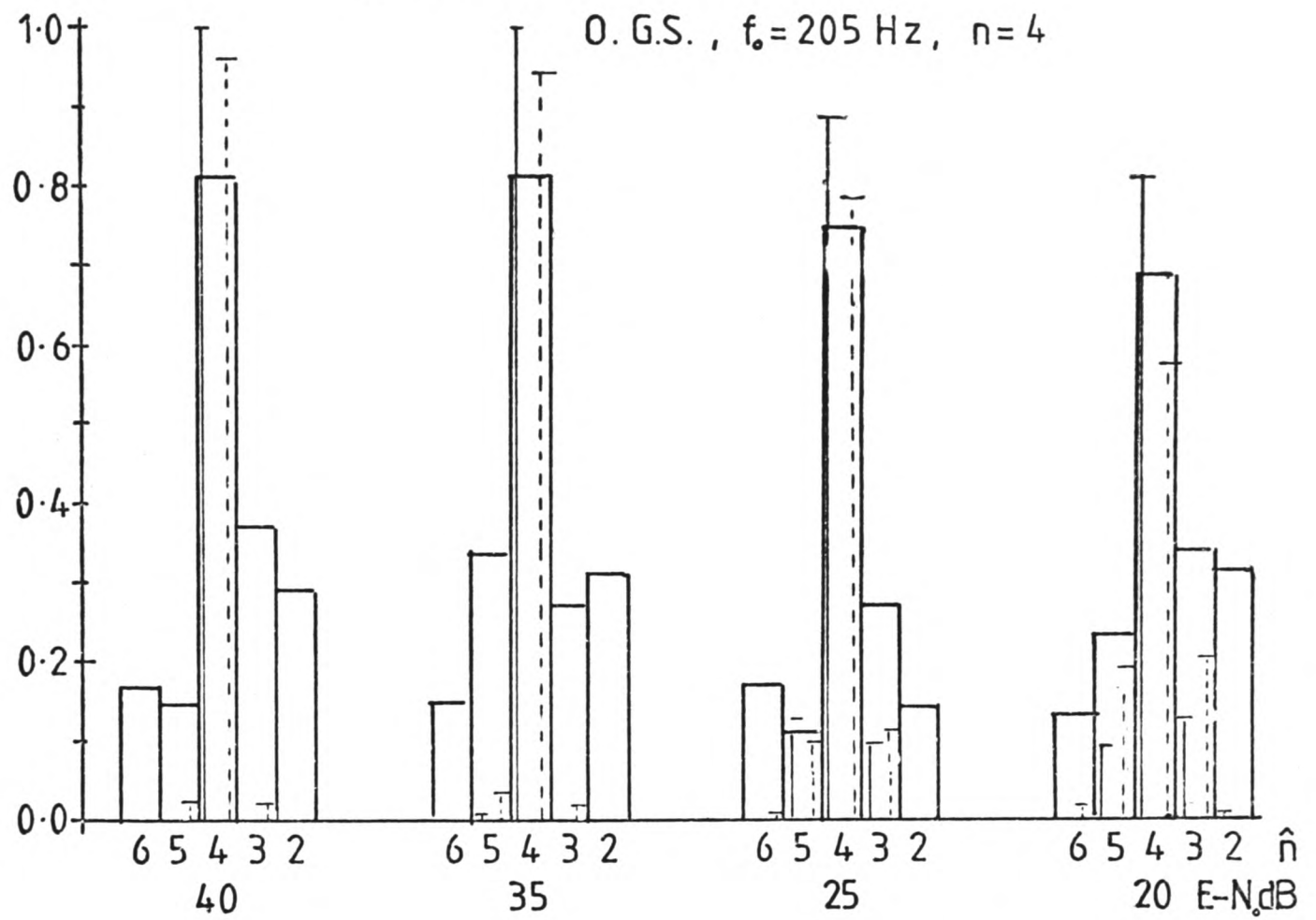
Fig. 10.8 a)



$p(\text{"same"})$

Fig. 10.8 b)

O.G.S., $f_0 = 205 \text{ Hz}$, $n = 4$



$n = 5$

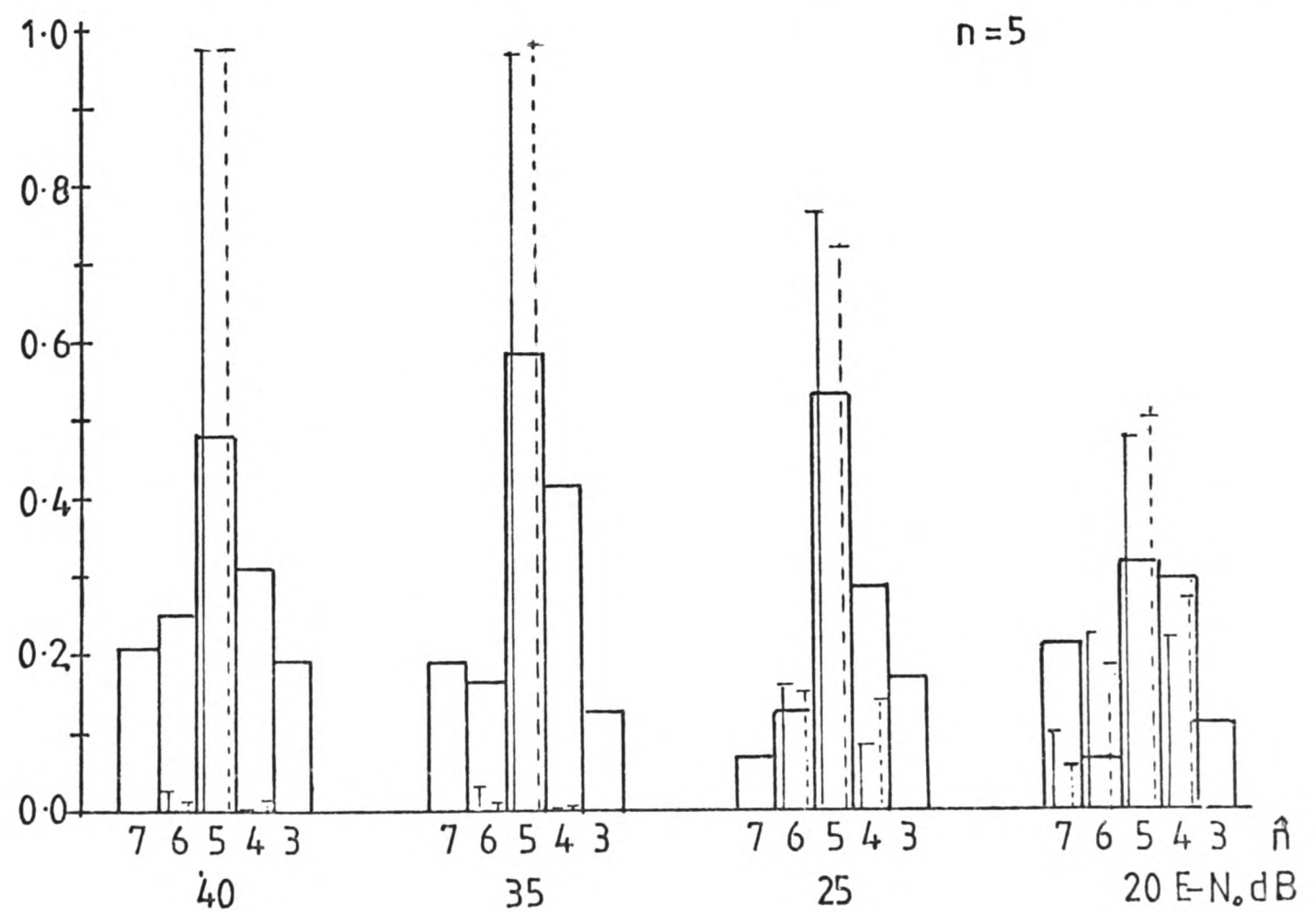


Figure 10.9. Observed proportions of "same" responses for Experiment 15a) for a standard signal fundamental frequency of 205 Hz only. Proportions of 40 trials. Panel a), observer A.F., panel b), observer G.S. The simulation results are also shown.

Fig. 10.9 a)

0. A.F. $f_0 = 205$ Hz, $n=4$

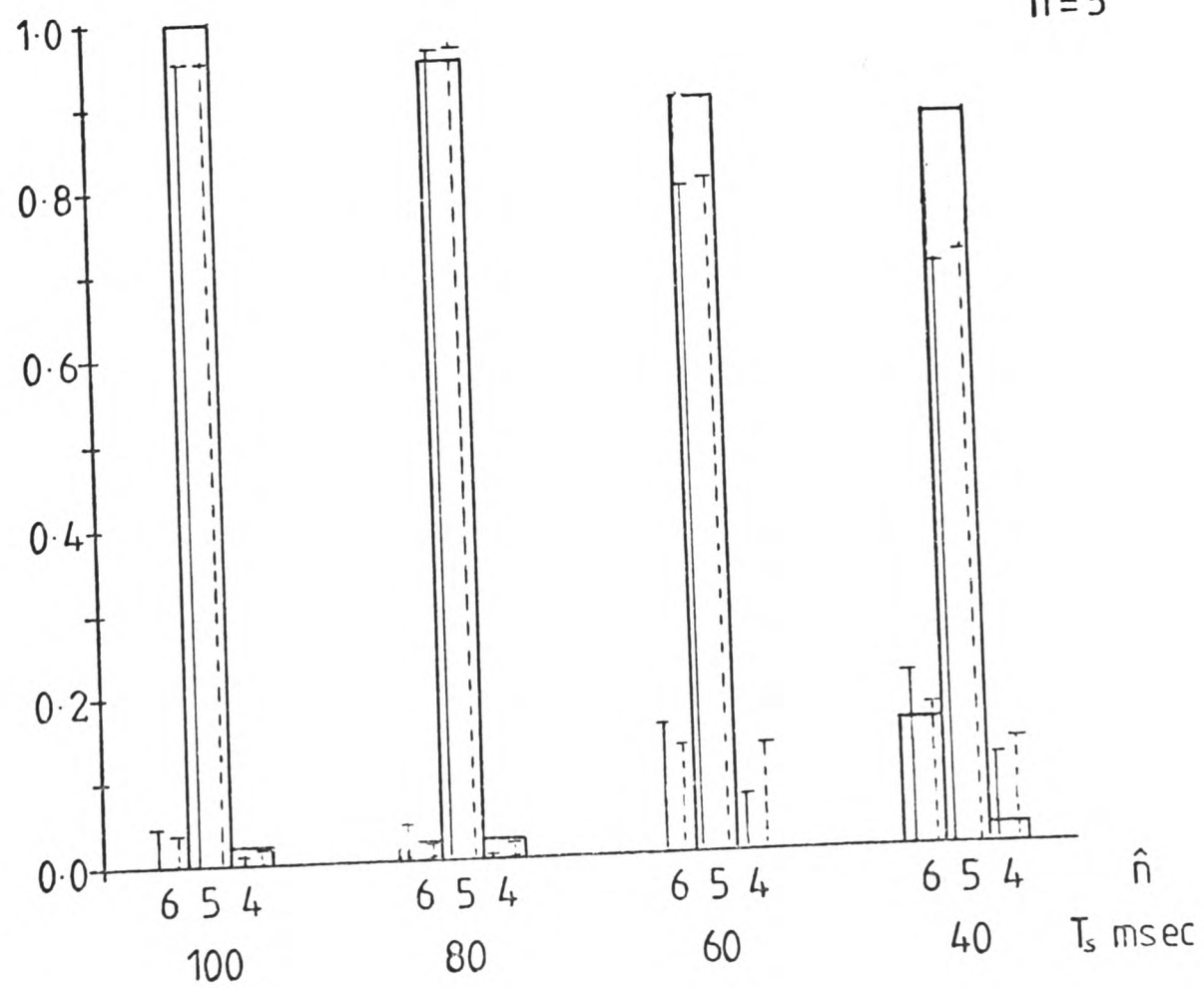
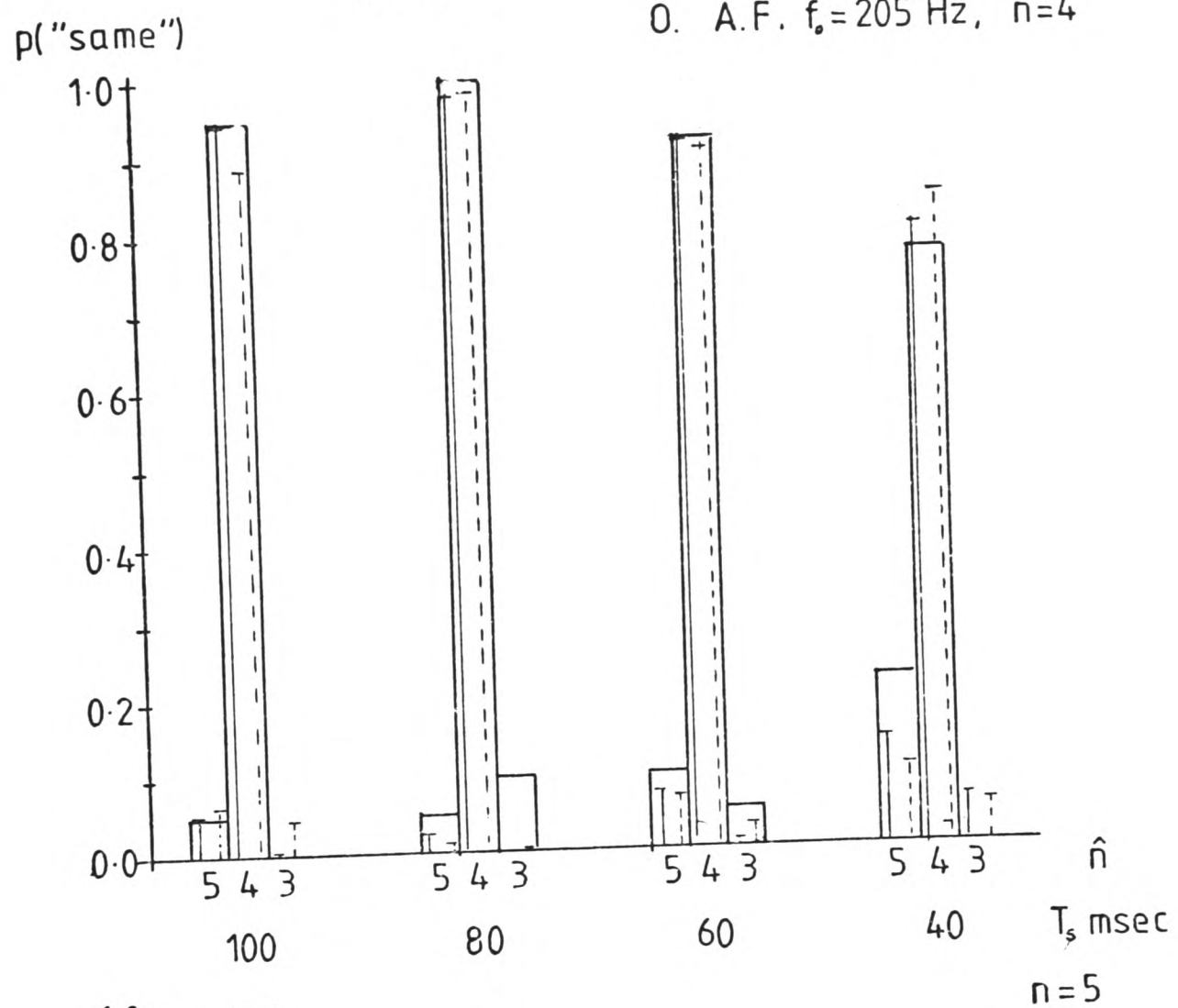


Fig. 10.9 b)

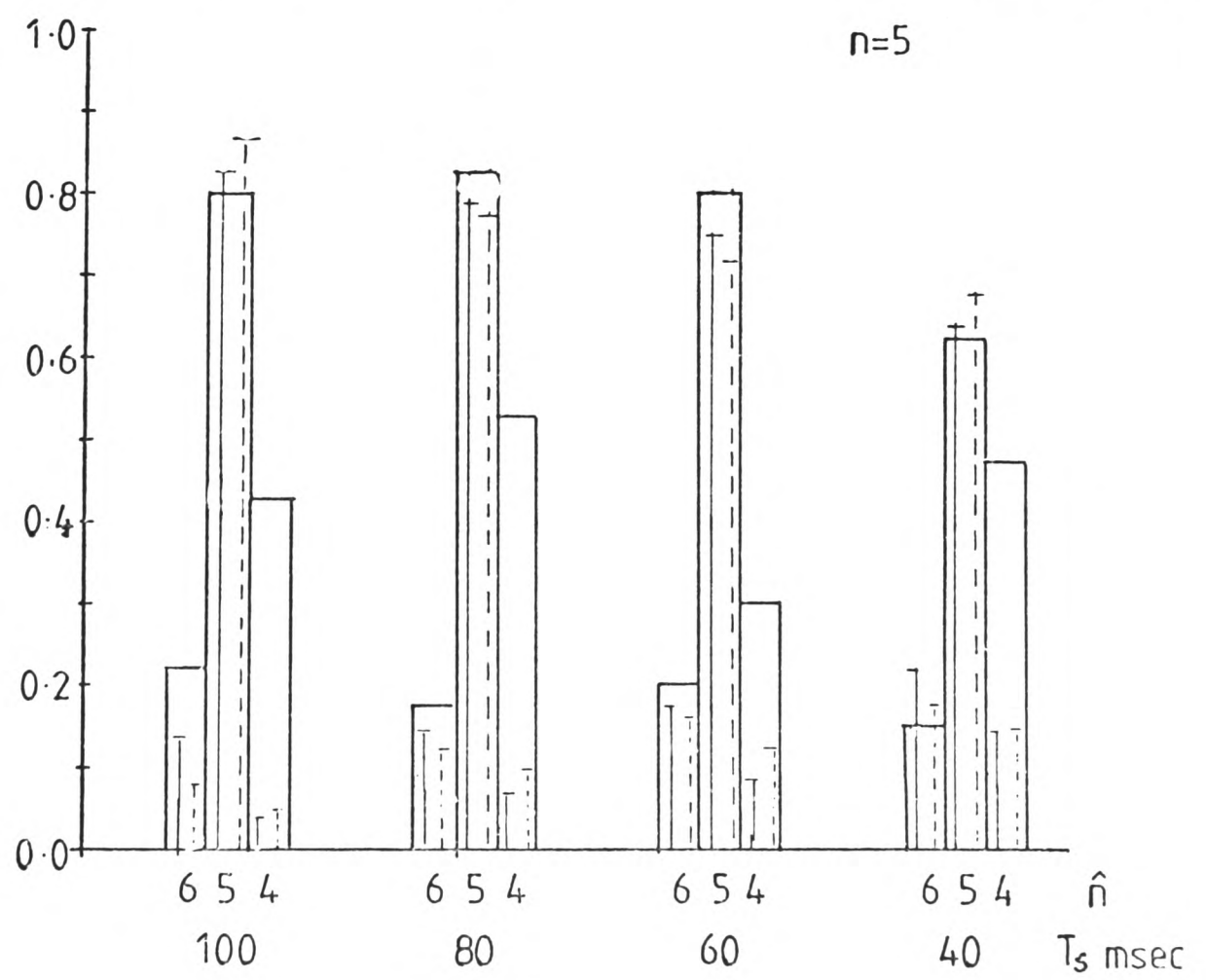
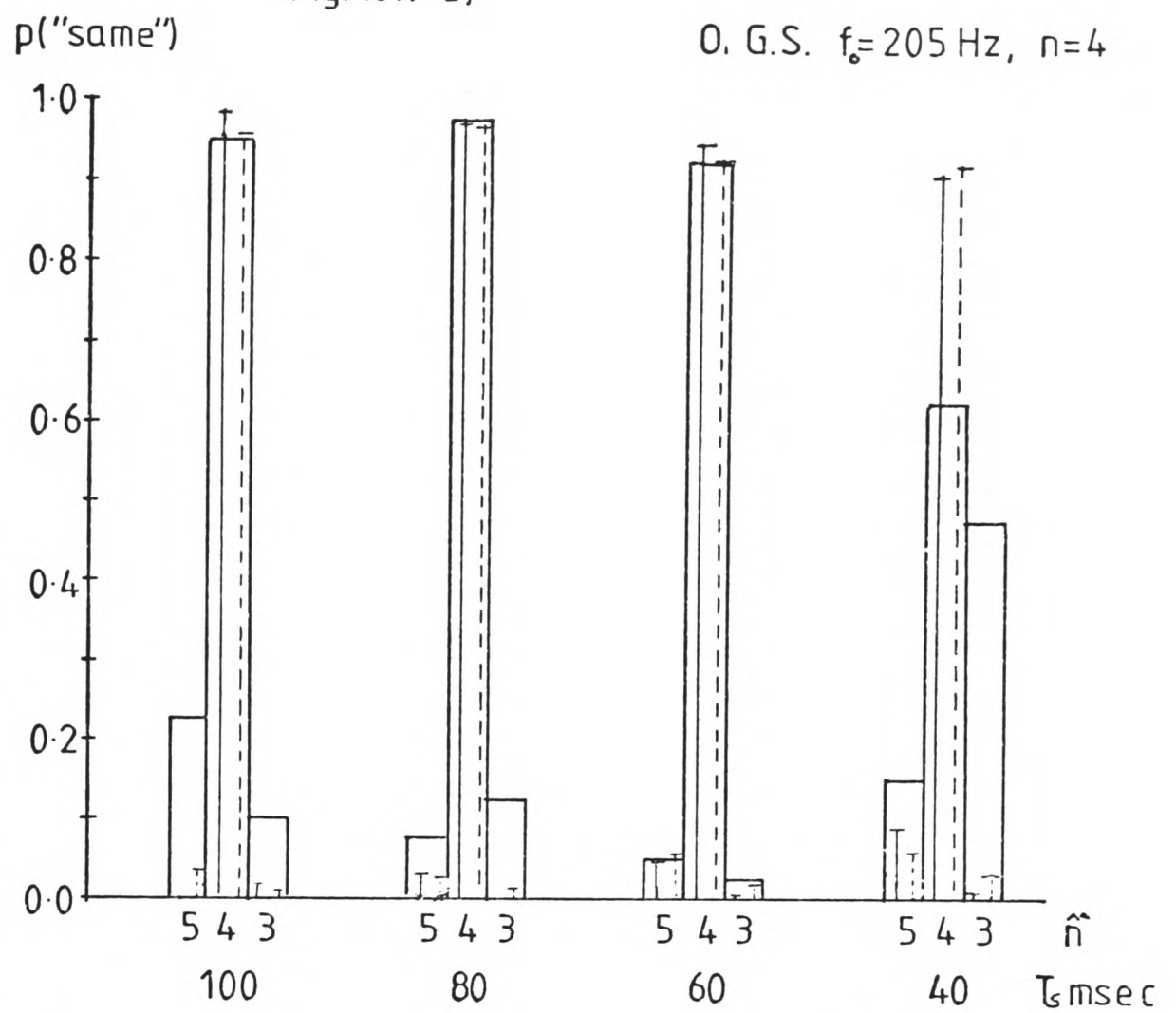


Figure 10.10. Observed proportions of "same" responses for Experiment 15b) for a standard signal fundamental frequency of 205 Hz only. Proportions of 40 trials. Panel a), observer A.F., panel b), observer G.S. The simulation results are also shown.

Fig.10.10a)

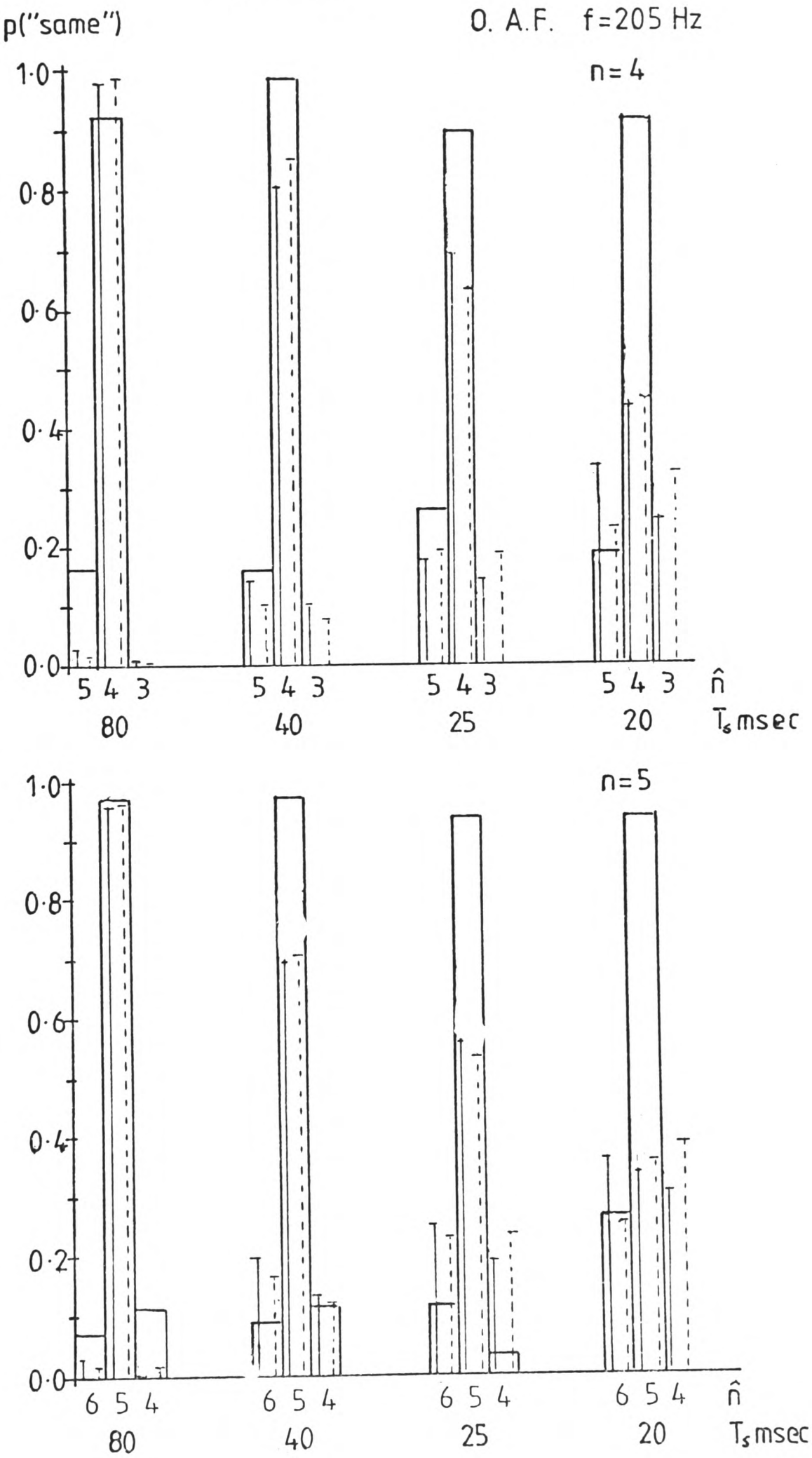
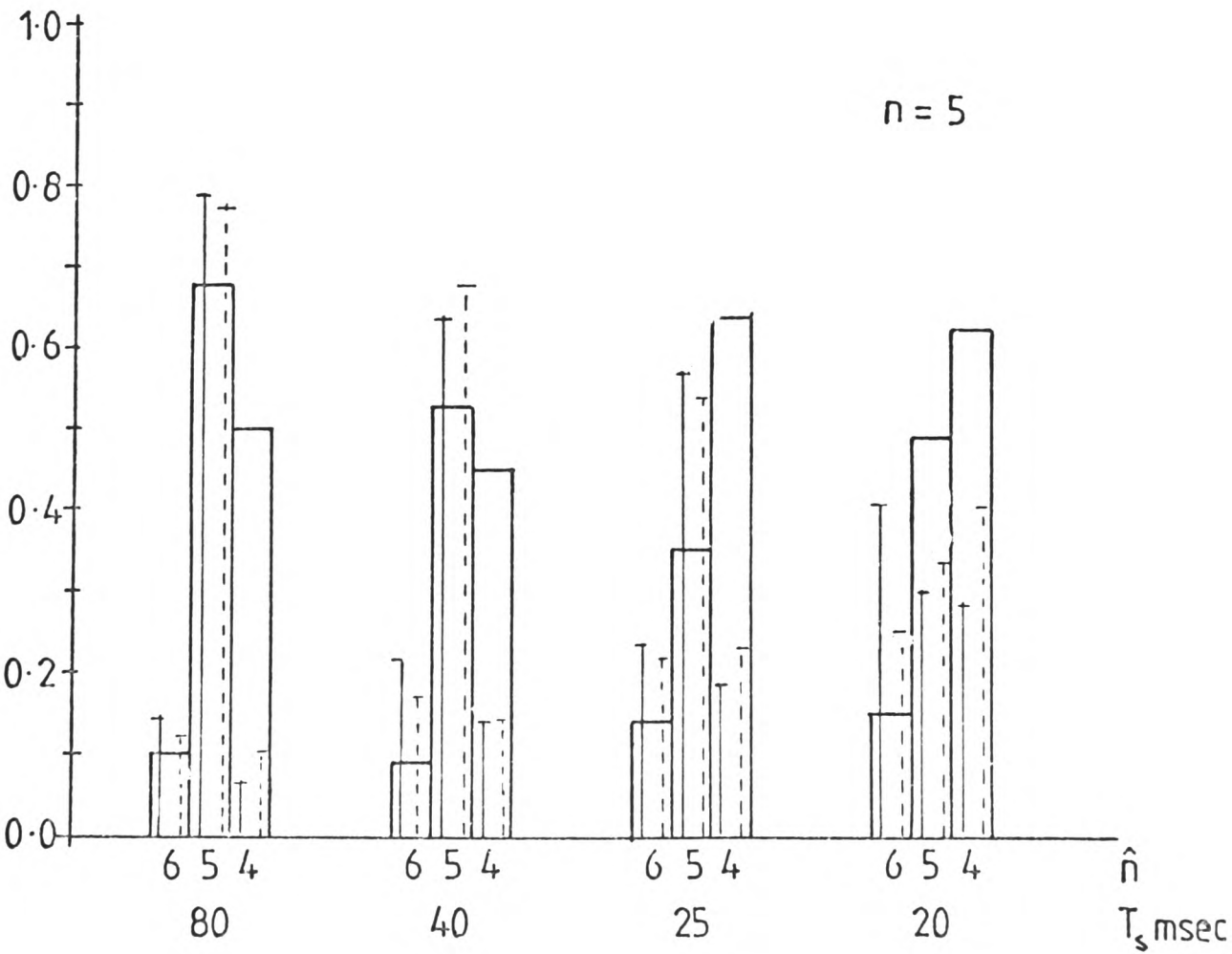
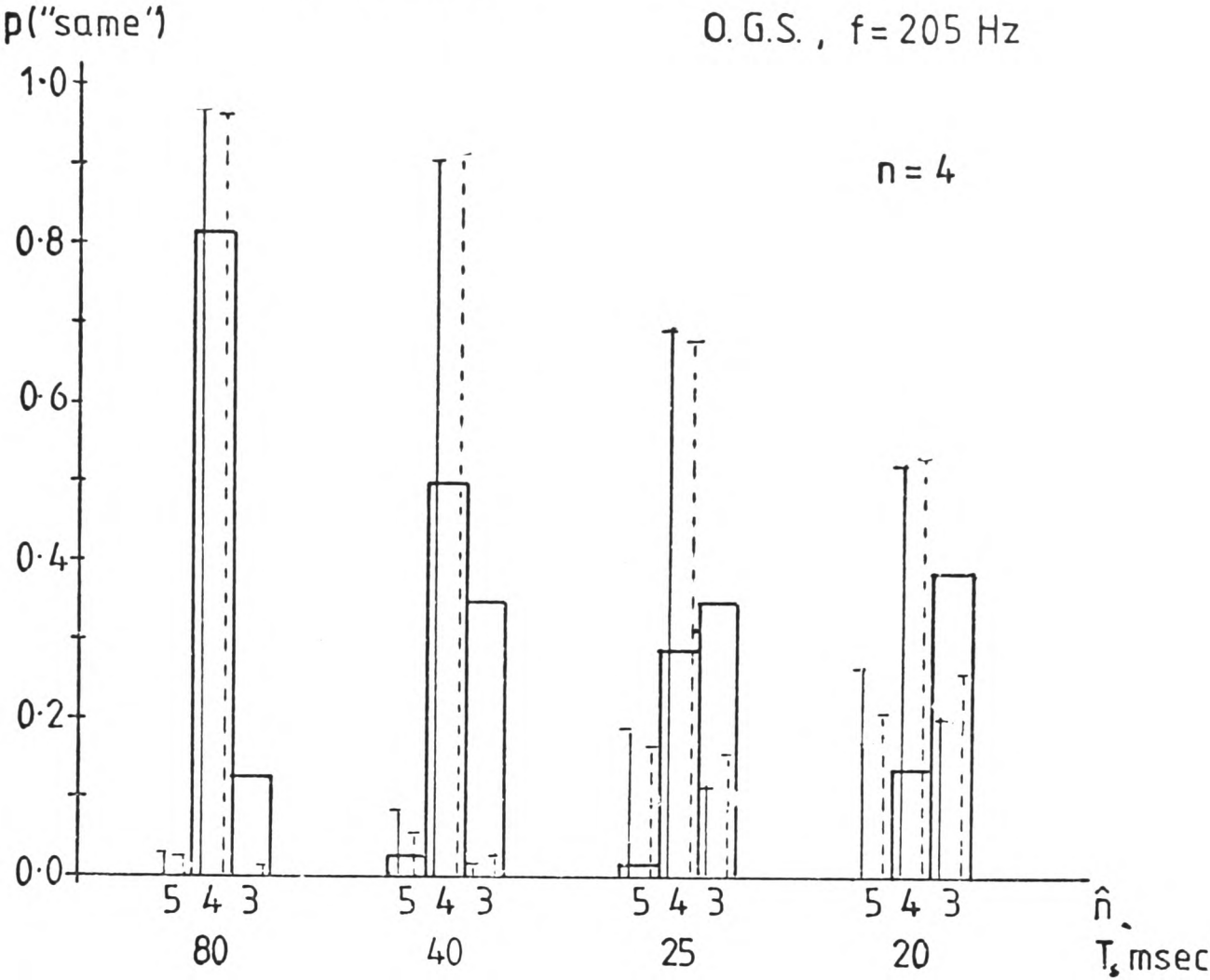


Fig. 10.10 b)

O.G.S., $f = 205 \text{ Hz}$



inter-observer differences of this sort.

As was noted above, observer A.F.'s identification performance at short signal durations was more accurate than that predicted by either model, and it was suggested that this discrepancy was in part due to this observer's prior knowledge of the stimuli. This suggestion receives some support from the distribution of "same" responses, for in Experiment 15b), this observer, unlike observer G.S., made virtually no "same" responses for a standard having $n=4$ and the comparison signal corresponding to an estimated harmonic number $\hat{n}=3$, while for standard durations of 25 msec and below, the two models predict that the pitch of the standard should correspond to pitch of this comparison signal with probabilities of between 0.10 and 0.20. Observer A.F. was aware the the stimulus set included no standard signals having $n=3$.

Summary and Conclusions.

The observed sensitivity of the two observers in the above three experiments is, in general, tolerably consistent with the predictions from computer simulations of the spectral-pattern recognition models of Goldstein (1973) and Wightman (1973b). One exception to this consistency was found for observer G.S. in the first of the three experiments, but in the two subsequent studies, after the observer had received more experience of the task, his performance approached the predicted levels. The second exception arose for observer A.F. at the shortest signal durations of Experiment 15 b). As was pointed out above, it is

likely that the predictions here were based upon overestimates of $\sigma(f'_j)/f_j$. A further factor which could account for this observer's better than predicted performance is that, as the experimenter, he was very familiar with the stimulus material, and was thus aware that the set of harmonic numbers represented by the standard signals was more limited than that set which the computer simulation considered. The distribution of 'same' responses for this observer are consistent with this suggestion.

CHAPTER 11

Introduction.

The studies of pitch perception reviewed in Chapter 1 can be best understood in terms of a spectral-pattern recognition theory, which assumes that the residue pitch of a tonal complex is derived from independent, noisy internal estimates of the component frequencies of the complex. The possibility that residual periodicity information, resulting from an incomplete auditory frequency analysis, may contribute to the perception of pitch cannot, however, be excluded.

The strongest evidence in favour of the spectral-pattern approach is that the prominence of residue pitch is greater the better the auditory resolution of component frequencies: Such a relationship is contrary to the predictions of a periodicity theory. This relationship is apparent in Ritsma's (1962) investigation of the existence region of residue pitch, in the dominance of low harmonic components (Plomp, 1976; Ritsma, 1967), in the limits to the audibility of residue pitch for two-component signals (Smootenburg, 1970; Sutton and Williams, 1970), and in Moore's (1973b) study of the limits to both component frequency identification and the audibility of residue

pitch. The only evidence from studies of tonal signals which favours a periodicity theory is that from studies of the effects on residue pitch of the relative phases of the components (Ritsma and Engel, 1964; Moore, 1977). Such effects are found only for rather poorly resolved components, and under conditions which also give rise to significant auditory distortion products (Buunen et al, 1974; Buunen and Bilsen, 1974, Smoorenburg, 1972a), which will cause phase-dependent changes in the effective spectrum of the signal. It has also been argued (Houtsma et al, 1980), that the pitch of amplitude-modulated noise is best explained by a periodicity theory, but that argument too ignores the effects of auditory non-linearities.

Even if periodicity information does make some contribution to the perception of residue pitch, it is clear that spectral information, whether coded spatially or temporally, is of the greater importance for the pitch of the majority of complex sounds. This realization has led to the formulation by Golstein (1973) and Wightman (1973b) of detailed theories of spectral-pattern processing. These two models are sufficiently formalised to allow the prediction, by computer simulation methods, of the dependence of a multimodal probability density function, representing the internal distribution of estimates of the residue pitch of a tonal complex, upon the standard deviations of a series of internal distributions representing estimates of the frequencies of the components of that complex. Both models proceed by seeking an estimate of fundamental frequency. Despite the considerably different approaches of the two models, their predictions are generally similar.

The key to an empirical investigation of these models is the estimation of the noisiness of component frequency estimation, $\sigma(f'_j)/f_j$. A number of investigators (Gerson and Goldstein, 1978; Goldstein, 1973; Goldstein et al, 1978; Houtsma, 1979) have made estimates of $\sigma(f'_j)/f_j$ from musical interval recognition, where the musical intervals are detectable as changes in residue pitch. Since the error rate in such a task is predicted by Goldstein's model to depend upon $\sigma(f'_j)/f_j$, this parameter can be indirectly estimated from performance. The resulting values of these estimates are between about 0.007 and 0.014 for components between 800 and 2000 Hz.

A more direct method of estimating this parameter is suggested by the theory of Signal Detection (Green and Swets, 1966), from which it may be shown that the relative frequency-difference limen (DL) is equivalent to the relative standard deviation of an internal representation of the frequency of a single component. Typical values of the relative frequency DL between 800 and 2000 Hz are between about 0.002 and 0.010 (Moore, 1973a; Weir, Jesteadt, and Green, 1977), and are too small to account for the error rate in musical interval recognition for residue pitch. Goldstein (1973) has accounted for this discrepancy by the assumption that $\sigma(f'_j)/f_j$ at pitch processing is largely determined by an internal noise process. This assumption also allows Goldstein's model to account for the residue pitch of inharmonic signals (de Boer, 1956; Patterson, 1973; Smoorenburg, 1970).

Discussion of the Experimental Results in Terms of Theories of Pitch Perception.

The pitch and frequency discrimination experiments reported above examine in detail the relation between component frequency discrimination and residue pitch discrimination for two-component harmonic complexes, and allow a comparison of discrimination performance within the same observers, under experimental conditions which remain as similar as possible for the two tasks. The estimates of $\sigma(f'_j)/f_j$ as a function of sensation level and signal-to-noise ratio from Experiments 3 and 7 suggest that, for components at 800 and 1000 Hz, the 4th and 5th harmonics of 200 Hz, this parameter may be as small as 0.002. Under comparable conditions of component frequency and level, Goldstein (1973) has estimated $\sigma(f'_j)/f_j$ at between 0.007 and 0.012.

Fundamental frequency discrimination between two signals composed of the 4th and 5th harmonics of fundamental frequencies close to 200 Hz was measured in Experiments 6 and 10. In accordance with the relationship between $\sigma(f'_0)/f_0$ and $\sigma(f'_j)/f_j$ predicted by the spectral-pattern models, the estimates of the former parameter were typically smaller than the minimum estimate of the latter. However, when fundamental frequency discrimination was measured in tasks requiring the comparison of pitches conveyed by different sets of harmonics of similar fundamental frequencies, (Experiments 5, 8, and 9), the resulting estimates of $\sigma(f'_0)/f_0$ were too large to be accounted for by the estimates of $\sigma(f'_j)/f_j$ from Experiments 3 and 7,

although the dependence of the estimates on presentation conditions was similar. The values of $\sigma(f'_j)/f_j$ consistent with the estimates of $\sigma(f'_0)/f_0$ from these experiments were closer to Goldstein's estimates from musical interval identification. This result can be interpreted as a confirmation of Goldstein's postulation of internal noise process in the channels conveying component frequency estimates to the pitch processor.

In order to reconcile the estimates of $\sigma(f'_0)/f_0$ from these two experimental tasks, it may be assumed that the internal noise postulated by Goldstein limits the minimum noisiness of the component frequency estimates at the pitch processor, but is effective only for fundamental frequency discrimination between signals having non-coincident components. Fundamental frequency discrimination for signals having coincident components may then be accounted for by a process which employs simultaneous component frequency comparisons rather than a fundamental frequency comparison.

The assumption of an added internal noise may seem to imply an inherently inefficient process. This impression, however, is false, for the noisy representation of component frequencies can also be interpreted as evidence that the pitch processor receives no more information than is necessary for the identification of pitch in the context of music, where mistunings of the order of 1% are inherent in the tempered scale and are generally unnoticed (Seashore, 1938; Ward, 1970).

Other explanations of the discrepancy between the two fundamental frequency discrimination tasks described here are also possible. The results do not rule out a periodicity theory explanation of residue pitch discrimination for signals with noncoincident components. An alternative explanation in terms of a spectral-pattern theory is that comparisons between pitches conveyed by non-coincident components are subject to a criterion variance which does not occur for comparisons of pitches conveyed by coincident components. It is also possible that some other difference between the experimental methods employed is responsible for the discrepancy, although no such cause is obvious.

The spectral-pattern theories, with $\sigma(f'_j)/f_j$ estimated with an internal noise component, and an allowance for frequency uncertainty, are also able to provide a fair account of the performance of the two observers in a 2IAX fundamental frequency identification task (Experiments 14 and 15), although performance here was also dependent upon the observers' familiarity with the task and the stimulus materials.

Conclusions with Respect to Theories of Pitch Perception.

A number of theories of pitch perception were described in Chapter 1. The evidence from the studies reviewed in that chapter and from the studies reported here does not favour Schouten's (1940) periodicity theory, at least as a general theory of all pitch phenomena, and the quantitative evidence from the above experiments is consistent with the predictions of

the spectral-pattern theories. The spectral-pattern models described by Wightman and by Goldstein are both consistent with the results described here. The only apparent distinction between the predictions of these two models is that Wightman's model, but not Goldstein's, predicts a direct dependence of residue pitch on the relative amplitudes of the signal components. Since Licklider's (1951, 1959) neural autocorrelation theory can be naturally extended to allow the use of information from well-resolved components, and is then very similar to Wightman's model, this theory is also consistent with the experimental results.

Future research on theories of pitch perception.

The spectral-pattern model described by Goldstein, and modified as described here, is amenable to further examination with the pitch discrimination methods employed here. The results obtained here do not allow a distinction between the assumptions that the internal noise implicated in the perception of residue pitch is due to a series of noisy frequency-selective channels, or to a criterion variance. However, these two assumptions may be tested. If the internal noise arises from frequency-selective channels, its contribution to the internal residue pitch pdf should be dependent on frequency uncertainty, since residue pitch discrimination is known to be subject to this effect (see Experiments 8 and 9). A criterion variance in the decision process should not be subject to this factor. If the internal noise is due to noisy frequency-selective channels, an experiment manipulating

frequency uncertainty should show comparable effects on pitch discrimination in both the residue pitch discrimination tasks described here. If a criterion variance is the source of the internal noise, the effects of frequency uncertainty in tasks requiring comparisons of the pitches conveyed by non-coincident components will be obscured by the criterion variance until the contribution from external and peripheral noise exceeds the criterion variance. This should not be the case for comparisons of pitches conveyed by coincident components, so that under the latter hypothesis, the effects of uncertainty on the two tasks would differ.

There is also a need for a detailed study of the effective spectra of signals such as amplitude-modulated noise, and the complex tonal signals showing relative phase effects, whose pitch appears to depend on periodicity information. It seems that these 'periodicity' effects may alternatively be due to spectral information derived from auditory non-linearities. There is, at present, insufficient data to distinguish between these hypotheses, particularly since estimates of the level of auditory distortion products are almost always made relative to the physical level of the primary components, so that little is known of the effects of auditory nonlinearities on the effective level of the primaries themselves.

It would also be valuable to examine in detail the effects on residue pitch of the relative amplitudes of the components of tonal complexes. The work of Ritsma (1967) suggests that relative amplitude is of little importance, and an investigation

of these effects would allow an empirical test between the predictions of Goldstein's and Wightman's models.

The Experimental Results and Theories of Frequency Discrimination.

The pitch and frequency discrimination data provide support for a Signal Detection model of frequency discrimination. The finding that relative fundamental frequency discrimination for signals with coincident components is determined by, and finer than, relative component frequency discrimination, is analagous to the finding that detection performance for complex signals is superior to detection for the components presented alone (Green, 1958; Green and Swets, 1966; Hall and Sondhi, 1977), and supports a multiple-band detection theory (Green and Swets, 1966). The data also show an effect of frequency uncertainty, which suggests that the observers were able to approach the optimal use of the available information.

The dependence of the discrimination data reported above on signal-to-noise ratio and signal duration is generally consistent with a version of the Energy Detection model of frequency discrimination (Green and Swets, 1966; Henning, 1967a), in which discrimination is ultimately limited by an internal noise with an equivalent spectral density of between 30 and 40 dB below the signal power. The results are also consistent with the duration dependence predicted by Goldstein and Srulovicz's (1977) temporal model of frequency discrimination, but this model does not make specific

predictions for the effects of signal-to-noise ratio.

Future Research on Theories of Frequency Discrimination.

The Energy Detection model of frequency discrimination was tested here by a comparison of the effects of signal-to-noise ratio and signal duration on fundamental frequency discrimination. A similar test with sinusoidal signals would be of value, as would an additional comparison of the effects of signal amplitude, where signal power and signal duration should have directly inverse effects on discrimination.

An experiment investigating the role of cues from transients in frequency discrimination for short duration signals would also be of value in the light of Moore's (1973a) contention that models such as the Energy Detector are unable to account for the limits to discrimination under these conditions.

Limits to the Generality of the Experimental Findings.

The experiments reported here are necessarily limited in their generality. In the first instance, most of the studies made use of two highly practiced observers, and there is no intention to claim that naive observers would show comparable performance. Secondly, the conclusions above are justifiable only for the small set of complex harmonic signals used for the experiments: This limitation arose partly from the limited existence region of the residue pitch of two-component signals (Houtsma and Goldstein, 1972; Smoorenburg, 1970; Sutton and

Williams, 1970), and partly from the requirement that the components of the complex signals should be well-resolved.

Other Applications of Spectral-Pattern Recognition.

Most meaningful auditory stimuli have a certain degree of spectral regularity, and the spectral-pattern theories could be usefully applied to the extraction of information related to such regularity. The pitch contours of speech can be extracted by a spectral-pattern recognition system based on Goldstein's model (Duifhuis, Willems, and Sluyter, 1978).

Spectral-pattern analysis may also play a part in the perception of the various spectral components of a number of simultaneous complex sounds: When a number of musical instruments play together, their respective frequency components will, with limits, be separated by auditory frequency analysis, yet the listener rarely has a fragmented perception of the sound of an individual instrument. Auditory localization and phase coherence may have some significance for our perceptions of such sounds, but in a reverberant environment, the reflection of sound and the interaction of direct and reflected sound waves will scramble much of this information for sources at any distance from the listener.

Probabilistic Pattern Perception.

The spectral-pattern recognition theories of residue pitch perception can be considered as examples of probabilistic pattern recognition in a more general sense. Complex harmonic

signals are ideal stimuli for such processes, because the perceptual features they contain are easily identified, and they possess an inherent spectral regularity, allowing the use of a universal template. There also exist other stimuli having similar properties, including spatially periodic complex visual patterns which have perceptual similarities to their absent fundamental spatial frequency component (Henning, Hertz, and Broadbent, 1975). The application of the logic of spectral-pattern recognition to such stimuli might increase our understanding of visual pattern perception. These stimuli are, however, somewhat exceptional among visual patterns in that their perceptual features can be identified with some confidence. The perceptual features of the majority of visual patterns, even those as simple as alphanumeric characters, are unknown. If hypothetical features can be identified, it may be possible to construct a probabilistic pattern recognition model for the use of these hypothetical features. Such a model could be tested by the application of Signal Detection theory to the discrimination and identification of the postulated features and of the whole patterns, in a manner analogous to the tests of the Spectral-pattern models of pitch perception described here.

- Bekesy, G. von. Experiments in hearing. McGraw-Hill, New York, 1963.
- Bilsen, F. A., and Ritsma, R. J. Repetition Pitch and its Implications for Hearing Theory. Acustica, 1970, 22, 63-73.
- Bilsen, F. A., and Goldstein, J. L. Pitch of Dichotically Delayed Noise and its Possible Spectral Basis. Journal of the Acoustical Society of America, 1974, 55, 292-296.
- Bilsen, F. A. Pitch of Noise Signals: Evidence for a "Central Spectrum". Journal of the Acoustical Society of America, 1977, 55, 292-296.
- Burns, E. M., and Viemeister, N. F. Nonspectral Pitch. Journal of the Acoustical Society of America, 1976, 60, 863-868.
- Buunen, T. J. F., Festen, J. M., Bilsen, F. A., and Brink, G. van den. Phase Effects in a Three Component Signal. Journal of the Acoustical Society of America, 1974, 55, 297-303.
- Buunen, T. J. F., and Bilsen, F. A. Subjective phase effects and combination tones. In E. Zwicker and E. Terhardt (Eds.). Facts and models in hearing. Springer, Berlin, 1974. pp 332-342.
- Boer, E. de. On the residue in hearing. Doctoral dissertation, University of Amsterdam, 1956.
- Boer, E. de. On the "residue" and auditory pitch perception. In W. D. Keidel and W. D. Neff (Eds.) Handbook of sensory physiology. Vol V.: Auditory system. Part 3: Clinical and special topics. Springer Verlag, Berlin, 1976. pp 479-593
- Boer, E. de. Pitch theories unified. In E. F. Evans, and J. P. Wilson, (Eds.). Psychophysics and physiology of hearing. Academic Press, London, 1977. pp 323-334.
- Creelman, C.D. Detection of Signals of Uncertain Frequency. Journal of the Acoustical Society of America, 1960, 32, 805-810.

- Creelman, C. D., and Macmillan, N. A. Auditory Phase and Frequency Discrimination: A Comparison of Nine Procedures. Journal of Experimental Psychology, Human Perception and Performance, 1979, 5, 146-156.
- Cudahy, E. A. Fundamental Pitch Discrimination for Monotic and Dichotic Two Tone Complexes. Journal of the Acoustical Society of America, 1975, 67 (S1), S20 (A).
- Duifhuis, H., Willems, L. F., and Sluyter, R. J. Pitch in Speech Institute for Perception Annual Progress Report, 13, 1978, 24-30
- Durlach, N. I. Binaural signal detection: Equalization and cancellation theory. In J. V. Tobias (Ed.). Foundations of modern auditory theory, Vol II. pp 369-462. Academic Press, New York, 1972
- Egan, J. P., and Hake, H. W. On the Masking Patterns of a Simple Auditory Stimulus. Journal of the Acoustical Society of America, 1950, 31, 1115-1120.
- Evans, E. F. Place and Time Coding of Frequency in the Peripheral Auditory System: Some Physiological Pros and Cons. Audiology, 1978, 17, 369-420.
- Evans, E. F., and Wilson, J. P. (Eds.). Psychophysics and physiology of hearing. Academic Press, London, 1977.
- Finney, D. J. Probit analysis. (3rd ed.). Cambridge University Press, Cambridge, 1971.
- Fletcher, H., and Munson, W. A. Loudness, its Definition, Measurement and Calculation. Journal of the Acoustical Society of America, 1933, 5, 82-108.
- Gässler, G. Über die Hörshwelle für Schallereignisse mit Verschieden Breitem Frequenzspektrum. Acustica, 1954, 4, 408-414.
- Gerson, A., and Goldstein, J. L. Evidence for a General Template

in Central Optimal Processing for Pitch of Complex Tones.

Journal of the Acoustical Society of America, 1978, 63, 498-510

Green, D. M. Detection of Multiple Component Signals in Noise.

Journal of the Acoustical Society of America, 1958, 30, 904-911.

Green, D. M. An introduction to hearing. Lawrence Earlbaum, New Jersey, 1976.

Green, D. M., McKey, M. J., and Licklider, J. C. R. Detection of a Pulsed Sinusoid in Noise as a Function of Frequency, Journal of the Acoustical Society of America, 1959, 31, 1446-1452.

Green, D. M., and Swets, J. A. Signal detection theory and psychophysics. Wiley, New York, 1966.

Goldstein, J. L. Aural combination tones. In R. Plomp and G. F. Smoorenburg (Eds.). Frequency analysis and periodicity detection in hearing. Sijthoff, Leiden, The Netherlands. 1970, pp 230-247

Goldstein, J. L. An Optimum Processor Theory for the Central Formation of the Pitch of Complex Tones. Journal of the Acoustical Society of America, 1973, 54, 1496-1516.

Goldstein, J. L. and Srulovicz, P. Auditory-Nerve spike intervals as an adequate basis for aural frequency measurement. In E. F. Evans and J. P. Wilson (Eds.). Psychophysics and physiology of hearing. Academic Press, London, 1977. pp 337-346.

Goldstein, J. L., Gerson, A., Srulovicz, P, and Furst, M. Verification of the Optimal Probabilistic Basis of Aural Processing in Pitch of Complex Tones. Journal of the Acoustical Society of America, 1978, 63, 486-497.

Hall, J. L., and Sondhi, M. M. Detection Threshold for a Two-Tone Complex. Journal of the Acoustical Society of America, 1977, 62, 636-640.

Hall, J. W. III., and Peters, R. W. Pitch of the Residue for

- Nonsimultaneous Harmonics. Journal of the Acoustical Society of America, 1980a, 67, S20 (A).
- Hall, J. W. III., and Peters, R. W. Pitch of the Residue Relative to Variables for Nonsimultaneous Harmonics. Journal of the Acoustical Society of America, 1980b, 68, S110 (A).
- Harris, G. G. Periodicity Detection by Using Gated Noise. Journal of the Acoustical Society of America, 1963, 35, 1229-1223.
- Harris, J. D. Pitch Discrimination. Journal of the Acoustical Society of America, 1952, 24, 750-755.
- Hays, W. L. Statistics. Holt, Rinehart, and Winston, New York, 1963.
- Helmholtz, H. L. F. von. Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. F. Veitweg & Sohn, Braunschweig, 1863. English translation of the 4th (1885) edition by A. J. Ellis, On the sensations of tone as a physiological basis for the theory of music. Reprinted by Dover, New York, 1954.
- Henning, G. B. Frequency Discrimination for Random Amplitude Tones. Journal of the Acoustical Society of America, 1966, 39, 336-339.
- Henning, G. B. A Model for Auditory Discrimination and Detection. Journal of the Acoustical Society of America, 1967a, 42, 1325-1334.
- Henning, G. B. Frequency Discrimination in Noise. Journal of the Acoustical Society of America, 1967b, 41, 774-777.
- Henning, G. B. A comparison of the effects of signal duration on frequency and amplitude discrimination. In R. Plomp and G. F. Smoorenburg (Eds.). Frequency analysis and periodicity detection in hearing. Sijthoff, Leiden, The Netherlands. 1970, pp 350-361.

- Henning, G. B., Hertz, B. G., and Broadbent, D. E. Some Experiments Bearing on the Hypothesis that the Visual System Analyses Spatial Patterns in Independent Bands of Spatial Frequency. Vision Research, 15, 1975, 887-897.
- Henning, G. B., and Grosberg, S. L. Effect of Harmonic Components on Frequency Discrimination. Journal of the Acoustical Society of America, 1968, 44, 1386-1389.
- Houtgast, T. Lateral suppression in hearing. Doctoral dissertation, Free University, Amsterdam, 1974.
- Houtgast, T. Subharmonic Pitches of a Pure Tone at Low S/N Ratio. Journal of the Acoustical Society of America, 1976, 60, 405-409.
- Houtsma, A. J. M. Musical Pitch of Two-Tone Complexes and Predictions by Modern Pitch Theories. Journal of the Acoustical Society of America, 1979, 66, 87-89.
- Houtsma, A. J. M., and Goldstein, J. L. The Central Origin of the Pitch of Complex Tones. Journal of the Acoustical Society of America, 1972, 51, 520-529.
- Houtsma, A. J. M., Wicke, R. W., and Ordubadi, A. Pitch of Amplitude-Modulated Noise and Predictions by Temporal and Spectral Theories. Journal of the Acoustical Society of America, 1980, 67, 1312-1322.
- Jesteadt, W., Weir, C. C., and Green, D. M. Intensity Discrimination as a Function of Frequency and Sensation Level. Journal of the Acoustical Society of America, 1977, 61, 169-177.
- Johnson, D. M., and Hafter, E. R. Uncertain-Frequency Detection: Cuing and Conditions of Observation. Perception and Psychophysics, 1980, 28, 143-149.
- Kaplan H. L., Macmillan, N. A., and Creelman, C. D. Tables of d' for Variable-Standard Discrimination Paradigms. Behavior

- Research Methods and Instrumentation, 1978, 10, 296-813.
- Keppel, G. Design and analysis. A researcher's handbook. Prentice-Hall, New Jersey, 1973
- Kiang, N. Y-S. Discharge of Single Fibers in the Cat's Auditory Nerve. Research Monograph 35, MIT Press, Cambridge, Massachusetts, 1965.
- Liang, C., and Chistovich, L. A. Frequency-Difference Limens as a Function of Tonal Duration. Soviet Physics Abstracts : Acoustics, 1961, 6, 75-80.
- Kendall, M. G. and Stuart, A. The advanced theory of statistics. Vol 2. Inference and relationship. Griffin, London, 1961.
- Licklider, J. C. R. A Duplex Theory of Pitch Perception. Experientia (Basel), 1951, 7/4, 128-134.
- Licklider, J. C. R. "Periodicity Pitch" and "Place Pitch". Journal of the Acoustical Society of America, 1954, 26, 945(A).
- Licklider, J. C. R. Three auditory theories. In S. Koch (Ed.). Psychology: A study of a science. Vol I. McGraw-Hill, New York, 1959, pp 41-144.
- Littler, T. S. The physics of the ear. Pergammon, Oxford, 1965.
- Luce, R. D. and Galanter, E. Discrimination. In R. D. Luce, R. R. Bush and E. Galanter (Eds.). Handbook of mathematical psychology. Vol I. Chap 4. Wiley, New York, 1963, pp 191-244.
- Luce, R. D., and Green, D. M. Neural Coding and Psychophysical Discrimination Data. Journal of the Acoustical Society of America, 1974, 56, 1554-1564.
- Macmillan, N. A., Kaplan, H. L., and Creelman, C. D. The Psychophysics of Categorical Perception. Psychological Review, 1977, 84, 452-471.
- Marill, T. M. Detection Theory and Psychophysics. Massachusetts

Institute of Technology: Research Laboratory of Electronics,
Technical Report No. 319, 1956.

Miller, G. A., and Taylor, W. G. The Perception of Repeated Bursts of Noise. Journal of the Acoustical Society of America, 1948, 20, 171-182.

Moore, B. C. J. Frequency Difference Limens for Short Duration Tones. Journal of the Acoustical Society of America, 1973a, 54, 610-619.

Moore, B. C. J. Some Experiments Relating to the Perception of Complex Tones. Quarterly Journal of Experimental Psychology, 1973b, 25, 451-475.

Moore, B. C. J. Effects of relative phase of the components on the pitch of three-component complex tones. In E. F. Evans and J. P. Wilson (Eds.). Psychophysics and physiology of hearing. Academic Press, London, 1977, pp 349-360.

Nelder, J. A., and Mead, R. A Simplex Method for Function Minimization. Computing Journal, 1965, 7, 308-313.

Papoulis, A. Probability, random variables and stochastic processes. McGraw-Hill, New York, 1965.

Patterson, R. D. Noise Masking of a Change in Residue Pitch. Journal of the Acoustical Society of America, 1969, 45, 1520-1524.

Patterson, R. D. The Effects of Relative Phase and the Number of Components on Residue Pitch. Journal of the Acoustical Society of America, 1973, 53, 1565-1572.

Patterson, R. D., and Johnson-Davies, D. Detection of a change in the pitch of AM noise. In E. F. Evans and J. P. Wilson (Eds.). Psychophysics and physiology of hearing. Academic Press, London, 1977, pp 364-371.

Patterson, R. D. and Nimmo-Smith, I. Off-Frequency Listening and

- Auditory Filter Asymmetry. Journal of the Acoustical Society of America, 1980, 67, 229-245.
- Patterson, R. D., and Wightman, F. L. Residue Pitch as a Function of Component Spacing. Journal of the Acoustical Society of America, 1973, 59, 1450-1459.
- Pierce, J. R., Lipes, R., and Cheetham, C. Uncertainty Concerning the Direct Use of Time Information in Hearing: Place Clues in White-Spectra Stimuli. Journal of the Acoustical Society of America, 1977, 61, 1609-1621.
- Pollack, I. Periodicity Pitch for Interrupted White Noise - Fact or Artifact. Journal of the Acoustical Society of America, 1969, 43, 316-323.
- Plomp, R. The Ear as a Frequency Analyser. Journal of the Acoustical Society of America, 1964, 36, 1628-1636.
- Plomp, R. Detectability threshold for combination tones. Journal of the Acoustical Society of America, 1965, 37, 1110-1123.
- Plomp, R. Experiments on tone perception. Academic thesis, Utrecht, 1966.
- Plomp, R. Pitch of complex tones. Journal of the Acoustical Society of America, 1967, 41, 1526-1533.
- Plomp, R. Aspects of tone sensation: A psychophysical study. Academic Press, London, 1976.
- Plomp, R., and Mimpen, A. M. The Ear as a Frequency Analyser. II. Journal of the Acoustical Society of America, 1968, 43, 764-767.
- Plomp, R., and Smoorenburg, G.F. (Eds.). Frequency analysis and periodicity detection in hearing. Sijthoff, Leiden, The Netherlands, 1970.
- Ritsma, R. J. Existence Region of the Tonal Residue. I. Journal of the Acoustical Society of America, 1962, 34, 1224-1229.

Ritsma, R. J. On Pitch Discrimination of Residue Tones.

International Audiology, 1963, 2, 34-37.

Ritsma, R. J. Frequencies Dominant in the Perception of the Pitch of Complex Sounds. Journal of the Acoustical Society of America, 1967, 42, 191-198.

Ritsma, R. J., and Engel, F. L. Pitch of Frequency-Modulated Signals. Journal of the Acoustical Society of America, 1964, 36, 1637-1644.

Robinson, D. W., and Dadson, R. S. A Redetermination of the Equal-Loudness Relations for Pure Tones. British Journal of Applied Physics, 1956, 7, 166-181.

Ronken, D. A. Some Effects of Bandwidth-Duration Constraints on Frequency Discrimination. Journal of the Acoustical Society of America, 1971, 49, 1232-1242.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. Phase-Locked Response to Low-Frequency Tones in Single Auditory Nerve Fibers of the Squirrel Monkey. Journal of Neurophysiology, 1967, 30, 769-793.

Rose, J. E., Hind, J. E., Anderson, D. J., and Brugge, J. F. Some Effects of Stimulus Intensity on Response of Auditory Nerve Fibers in the Squirrel Monkey. Journal of Neurophysiology, 1971, 34, 685-699.

Seashore, C. E. Psychology of music. McGraw-Hill, New York, 1938.

Schafer, T. H. and Gales, R. S. Auditory Masking of Multiple Tones by Random Noise. Journal of the Acoustical Society of America, 1949, 21, 392-398.

Scharf, B. Critical bands. In J. V. Tobias (Ed.). Foundations of modern auditory theory, Vol I. Academic Press, New York. 1970, pp 157-202.

- Schodder, G. F., and David, E. E. Jr. Pitch Discrimination of Two-Frequency Complexes. Journal of the Acoustical Society of America, 1960, 32, 1426-1435.
- Schouten, J. F. The Perception of Subjective Tones. Proceedings. Koninklijke Nederlandse Akademie van Wetenschappen, 1938, 41, 1086-1093.
- Schouten, J. F. The Residue and the Mechanism of Hearing. Proceedings. Koninklijke Nederlandse Akademie van Wetenschappen, 1940, 43, 991-999.
- Schouten, J. F. The residue revisited. In R. Plomp and G. F. Smoorenburg (Eds.). Frequency analysis and periodicity detection in hearing. Sijthoff, Leiden, The Netherlands. 1970, pp 41-54.
- Schouten, J. F., Ritsma, R. J., and Cardozo, B. L. Pitch of the Residue. Journal of the Acoustical Society of America, 1962, 34, 1418-1424.
- Shannon, C. E., and Weaver, W. The mathematical theory of communication. University of Illinois Press, Urbana, 1949.
- Shower, E. G., and Biddulph, R. Differential Pitch Sensitivity of the Ear. Journal of the Acoustical Society of America, 1931, 3, 275-287.
- Siebert, W. M. Frequency Discrimination in the Auditory System: Place or Periodicity Mechanisms? Proceedings of the IEEE, 1970, 58, 723-730.
- Small, A. M., and Campbell, R. A. Masking of Pulsed Tones by Bands of Noise. Journal of the Acoustical Society of America, 1961, 33, 1570-1576.
- Smoorenburg, G. F. Pitch Perception of Two Frequency Stimuli. Journal of the Acoustical Society of America, 1970, 48, 924-942.
- Smoorenburg, G. F. Audibility Region of Combination Tones. Journal

- of the Acoustical Society of America, 1972a, 52, 603-614.
- Smootenburg, G. F. Combination Tones and Their Origin. Journal of the Acoustical Society of America, 1972b, 52, 615-632.
- Smootenburg, G. F. On the mechanisms of combination tone generation and lateral inhibition in hearing. In E. Zwicker and E. Terhardt (Eds.). Facts and models in hearing. Springer, Berlin, 1974. pp 332-342.
- Stevens, S. S. and Newman, E. B. The Localization of Actual Sources of Sound. American Journal of Psychology, 1936, 48, 297-306.
- Sutton, R. A., and Williams, R. P. Residue Pitches from Two-Tone Complexes. Journal of Sound and Vibration, 1970, 13, 195-199.
- Tanner, W. P., Jr., Swets, J. A., and Green, D. M., Some General Properties of the Hearing Mechanism. University of Michigan: Electronic Defense Group, Technical Report No. 30, 1956.
- Terhardt, E. Pitch, Consonance, and Harmony. Journal of the Acoustical Society of America, 1974, 55, 1061-1069.
- Thurlow, W. R. Perception of Low Auditory Pitch: A Multicue, Mediation Theory. Psychological Review, 1963, 70, 461-470.
- Turnbull, W. W. Pitch Discrimination as a Function of Tonal Duration. Journal of Experimental Psychology, 1944, 34, 302-316.
- Veniar, F.A. Signal Detection as a Function of Frequency Ensemble, I. Journal of the Acoustical Society of America, 1958a, 30, 1020-1024.
- Veniar, F.A. Signal Detection as a Function of Frequency Ensemble, II. Journal of the Acoustical Society of America, 1958b, 30, 1075-1078.
- Ward, W. D. Musical perception. In J. V. Tobias (Ed.). Foundations of modern auditory theory, Vol I. Academic Press, New York, 1970.
- Walliser, K. von. Über ein Funktionsschema für die Bildung der Periodentonhöhe aus dem Schallreiz. Kybernetik, 1969a, 6, 65-72.
- Walliser, K. von. Zur Unterschiedsschwelle der Periodentonhöhe.

Acustica, 1969b, 21, 329-336.

Weber, D. L. Growth of Masking and the Auditory Filter. Journal of the Acoustical Society of America, 1977, 62, 424-429.

Wegel, R. L., and Lane, C. E. The Auditory Masking of One Pure Tone by Another, and its Probable Relation to the Dynamics of the Inner Ear. Physics Review, 1924, 23, 266-285.

Weir, C. C., Green, D. M., Hafter, E. R. and Burkhardt, S. Detection of a Toneburst in Continuous- and Gated-noise Maskers; Defects of Signal Frequency, Duration, and Masker Level. Journal of the Acoustical Society of America, 1977b, 61, 1298-1300

Weir, C. C., Jesteadt, W., and Green, D. M. Frequency Discrimination as a Function of Frequency and Sensation Level. Journal of the Acoustical Society of America, 1977, 61, 178-184.

Whitfield, I. C. Central nervous processing in relation to spatio-temporal discrimination of auditory patterns. In R. Plomp and G. F. Smoorenburg (Eds.). Frequency analysis and periodicity detection in hearing. Sijtoff, Leiden, The Netherlands, 1970.

Whitfield, I. C. The neural code. In E. C. Carterette and M. P. Friedman (Eds.). Handbook of perception. Vol IV. Hearing. Academic Press, New York, 1978.

Wiener, N. Cybernetics. Wiley, New York, 1948

Wightman, F. L. Pitch and Stimulus Fine Structure. Journal of the Acoustical Society of America, 1973a, 54, 397-406.

Wightman, F. L. The Pattern-Transformation Model of Pitch. Journal of the Acoustical Society of America, 1973b, 54, 407-418.

Reference Notes.

1. Moore, B. C. J. Personal communication, 1980.
2. N. A. G. Algol 60 Library Manual. Mark 7. National Algorithm Group Ltd., Oxford, 1978.
3. Algol (E): Language. International Computers Ltd. Technical Publication 6855. I.C.L., London, 1976.
4. The Rogalgot Algol-60 System for the PDP-8. RHA (Minisystems) Ltd., Oxford.
5. Cudahy, E. A. Manuscript of paper presented at the Acoustical Society of America Meeting, April 9th, 1975.

APPENDIX A

Results of simulation procedure and experimental data for Experiments 14 and 15.

The legend \underline{n} represents the physical lower harmonic number of the standard signal, and the legend \hat{n} represents the estimated lower harmonic number as in Goldstein's model. The obtained pitch modes from the simulation of Wightman's model correspond in frequency to those for Goldstein's model. The pitch mode at $\hat{n} = \underline{n}$ has a mean at or close to \underline{f}_0 ; modes for $\hat{n} > \underline{n}$ represent pitches below \underline{f}_0 , and modes for $\hat{n} < \underline{n}$ pitches above \underline{f}_0 .

For the simulation results, the value tabulated under each mode is the proportion of 400 trials on which the simulation gave a best estimate of fundamental frequency from the \hat{n} th mode. For the observed data, the entry is the proportion of trials on which the observer made the response "same" to the indicated standard and comparison signal pair. The hit rate (H) is the proportion at $\hat{n} = \underline{n}$, and the false alarm rate (FA) the mean proportion over the other modes. The numerical entry to the right of the table is the fundamental frequency of the standard signal. The maximum value of $\underline{d'}$ given by Kaplan et al (1977) is 6.93, for a hit rate of 0.99, and a false alarm rate of 0.01.

The number of trials per row is 400 for all the simulations.
The number of trials per cell for observed data is given in the table below.

Experiment 14. Observer A.F. 48 trials per cell for each standard fundamental frequency, and 144 trials per cell for the mean data.							
E-N _o n	6	5	4 (H)	3	2	FA	d'
20dB 4							Predicted
	0.0000	0.1900	0.7200	0.0900	0.0000	0.0700	3.61 Goldstein
	0.0300	0.1750	0.5300	0.2650	0.0000	0.1180	2.65 Wightman
							Observed
	0.0000	0.2292	0.5625	0.0833	0.0000	0.0781	3.06 $f_o = 205$
	0.0208	0.0000	1.0000	0.0417	0.0000	0.0156	>6.55 170
	0.3542	0.1042	0.8333	0.1667	0.0000	0.1563	3.35 258
	0.1250	0.1111	0.7986	0.0972	0.0000	0.0833	3.80 mean
							Predicted
	0.0000	0.0875	0.8950	0.0175	0.0000	0.0263	4.99 Goldstein
25dB 4	0.0000	0.1350	0.7200	0.1450	0.0000	0.0700	3.61 Wightman
							Observed
	0.0000	0.1250	0.9792	0.0000	0.0208	0.0365	5.77 $f_o = 205$
	0.0000	0.0208	0.9792	0.0000	0.0000	0.0052	6.58 170
	0.2083	0.0208	0.9375	0.0833	0.0000	0.0781	4.65 258
	0.0694	0.0556	0.9653	0.0278	0.0069	0.0399	5.54 mean
							Predicted
	0.0000	0.0525	0.9425	0.0050	0.0000	0.0144	5.95 Goldstein
	0.0000	0.0675	0.8850	0.0475	0.0000	0.0288	4.92 Wightman
							Observed
30dB 4	0.0208	0.0833	0.9583	0.0000	0.0000	0.0260	5.56 $f_o = 205$
	0.0000	0.0417	1.0000	0.0000	0.0000	0.0104	>6.93 170
	0.0833	0.0000	1.0000	0.1250	0.0000	0.0521	>5.97 258
	0.0347	0.0417	0.9861	0.0417	0.0000	0.0295	6.30 mean
							Predicted
	0.0000	0.0625	0.9350	0.0025	0.0000	0.0163	5.56 Goldstein
	0.0000	0.0850	0.8500	0.0650	0.0000	0.0375	4.51 Wightman
							Observed
	0.0000	0.0417	0.9583	0.0208	0.0000	0.0156	5.81 $f_o = 205$
	0.0000	0.0208	1.0000	0.0000	0.0000	0.0052	>6.93 170
35dB 4	0.1250	0.0000	1.0000	0.0625	0.0000	0.0469	>5.97 258
	0.0417	0.0208	0.9861	0.0278	0.0000	0.0226	6.55 mean

Experiment 14 - Observer A.F. - Continued.							
		\hat{n}					
		7	6	5 (H)	4	3	FA
20dB	5						
							Predicted
		0.0100	0.1800	0.6675	0.1425	0.0000	0.0083 3.36 Goldstein
		0.0050	0.1925	0.6125	0.1900	0.0000	0.0969 3.02 Wightman
							Observed
		0.0208	0.2292	0.8333	0.1042	0.0625	0.1042 3.75 $f_o = 205$
		0.0625	0.0208	0.8125	0.2083	0.0000	0.0729 3.94 170
		0.0000	0.0000	0.9167	0.0417	0.1250	0.0417 4.95 258
		0.0278	0.0833	0.8542	0.1181	0.0625	0.0729 4.12 mean
							Predicted
25dB	5						
							Predicted
		0.0000	0.0575	0.9275	0.0150	0.0000	0.0180 5.85 Goldstein
		0.0000	0.0500	0.9275	0.0225	0.0000	0.0181 5.47 Wightman
							Observed
		0.0000	0.0625	0.9583	0.1875	0.0000	0.0625 5.10 $f_o = 205$
		0.0000	0.0000	1.0000	0.1042	0.0000	0.0260 >6.30 170
		0.0208	0.0000	0.9375	0.0208	0.0625	0.0260 5.32 258
		0.0069	0.0208	0.9653	0.1042	0.0208	0.0382 5.54 mean
							Predicted
30dB	5						
							Predicted
		0.0000	0.0475	0.9450	0.0075	0.0000	0.0138 6.06 Goldstein
		0.0000	0.0375	0.9450	0.0175	0.0000	0.0138 6.06 Wightman
							Observed
		0.0208	0.0625	0.9792	0.1667	0.0000	0.0625 5.49 $f_o = 205$
		0.0000	0.0000	0.9792	0.0417	0.0000	0.0104 6.58 170
		0.0417	0.0208	0.9583	0.0000	0.0625	0.0313 5.56 258
		0.0208	0.0278	0.9722	0.0694	0.0208	0.0347 5.73 mean
							Predicted
35dB	5						
							Predicted
		0.0000	0.0600	0.9400	0.0000	0.0000	0.0150 5.75 Goldstein
		0.0000	0.0300	0.9500	0.0200	0.0000	0.0125 6.06 Wightman
							Observed
		0.0000	0.0417	0.9375	0.0625	0.0000	0.0260 5.32 $f_o = 205$
		0.0000	0.0000	0.9792	0.0417	0.0208	0.0156 6.19 170
		0.0417	0.0000	0.9792	0.0000	0.0417	0.0208 6.19 258
		0.0139	0.0139	0.9653	0.0347	0.0208	0.0208 5.97 mean

Experiment 14 - Predictions and Results.

Observer G.S. 48 trials per cell for each standard fundamental frequency, and 144 trials per cell for the mean data, except at E-N _o of 25 dB, where there were 47 and 141 trials.							
\hat{n}							
E-N _o n	6	5	4 (H)	3	2	FA	d'
20dB 4							Predicted
	0.0000	0.0525	0.8175	0.1275	0.0025	0.0456	4.22 Goldstein
	0.0200	0.1950	0.5825	0.2025	0.0000	0.1044	2.94 Wightman
							Observed
	0.1250	0.2292	0.6875	0.3333	0.3125	0.2500	2.37 f _o =205
	0.1042	0.1250	0.6458	0.3750	0.2083	0.2031	2.49 170
25dB 4	0.0833	0.1250	0.2708	0.2500	0.1042	0.1406	1.65 258
	0.1042	0.1597	0.5347	0.3194	0.2083	0.1979	2.15 mean
							Predicted
	0.0000	0.0125	0.8925	0.0950	0.0000	0.0269	4.92 Goldstein
	0.0025	0.0950	0.7900	0.1125	0.0000	0.0525	4.10 Wightman
							Observed
35dB 4	0.1721	0.1069	0.7482	0.2718	0.1379	0.1721	2.97 f _o =205
	0.2165	0.1069	0.7020	0.4212	0.5392	0.3209	2.10 170
	0.0634	0.2563	0.2772	0.2545	0.1042	0.1696	1.44 258
	0.1507	0.1567	0.5758	0.3158	0.2604	0.2209	2.19 mean
							Predicted
	0.0000	0.0050	0.9950	0.0000	0.0000	0.0013	>6.93 Goldstein
40dB 4	0.0000	0.0325	0.9475	0.0200	0.0000	0.0131	6.06 Wightman
							Observed
	0.1458	0.3333	0.8125	0.2708	0.3125	0.2656	2.72 f _o =205
	0.1250	0.1458	0.7292	0.1667	0.1042	0.1354	3.08 170
	0.2083	0.2083	0.2708	0.3750	0.1250	0.2292	0.82 258
	0.1597	0.2292	0.6042	0.2708	0.1806	0.2101	2.30 mean
40dB 4							Predicted
	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	>6.93 Goldstein
	0.0000	0.0200	0.9625	0.0175	0.0000	0.0094	6.19 Wightman
							Observed
	0.1667	0.1458	0.8125	0.3750	0.2917	0.2448	2.85 f _o =205
	0.1115	0.1667	0.7500	0.3333	0.1667	0.1945	2.86 170
	0.1667	0.2292	0.2083	0.2500	0.0490	0.1737	0.93 258
	0.1483	0.1806	0.5903	0.3194	0.1691	0.2043	2.32 mean

Experiment 14 - Observer G.S. - Continued.

		\hat{n}					FA	d'	
E-No.	n	7	6	5 (H)	4	3			
20dB	5	0.0925	0.2175	0.4725	0.2175	0.0000	0.1319	2.42	Predicted Goldstein
		0.0525	0.1800	0.5000	0.2675	0.0000	0.1250	2.50	Wightman
	5	0.2083	0.0625	0.3125	0.2917	0.1042	0.1667	1.59	Observed $f_o = 205$
		0.0833	0.2083	0.2917	0.5000	0.3750	0.2917	0.00	170
		0.0208	0.1042	0.1250	0.0833	0.1250	0.0833	1.40	258
		0.1042	0.1250	0.2431	0.2917	0.2014	0.1806	1.09	mean
		0.0000	0.1600	0.7600	0.0800	0.0000	0.0600	3.86	Predicted Goldstein
		0.0000	0.1475	0.7175	0.1350	0.0000	0.0706	3.61	Wightman
		0.0652	0.1268	0.5317	0.2971	0.1667	0.1640	2.38	Observed $f_o = 205$
		0.0435	0.1286	0.3007	0.3822	0.3161	0.2176	1.14	170
		0.0426	0.0417	0.1042	0.0634	0.1477	0.0738	1.20	258
		0.0504	0.0990	0.3122	0.2476	0.2101	0.1518	1.75	mean
25dB	5	0.0000	0.0300	0.9675	0.0025	0.0000	0.0081	6.36	Predicted Goldstein
		0.0000	0.0125	0.9825	0.0050	0.0000	0.0044	6.58	Wightman
	5	0.1875	0.1667	0.5833	0.2917	0.1250	0.1927	2.35	Observed $f_o = 205$
		0.1042	0.1250	0.3125	0.4167	0.2500	0.2240	1.20	170
		0.0417	0.0417	0.1042	0.0417	0.1458	0.0677	1.20	258
		0.1111	0.1111	0.3333	0.2500	0.1736	0.1615	1.75	mean
		0.0000	0.0250	0.9725	0.0025	0.0000	0.0069	6.36	Predicted Goldstein
		0.0000	0.0125	0.9750	0.0125	0.0000	0.0063	6.58	Wightman
		0.2083	0.2500	0.4792	0.3125	0.1875	0.2396	1.78	Observed $f_o = 205$
		0.1042	0.0625	0.3125	0.3333	0.2708	0.1927	1.44	170
		0.0625	0.0000	0.0417	0.0833	0.0417	0.0469	-0.95	258
		0.1250	0.1042	0.2778	0.2431	0.1667	0.1597	1.53	mean

Experiment 15 - Predictions and Results.

The letter 'a' or 'b' after the standard fundamental frequency refers to the presentation order of the standard and comparison signal. Order a was with the comparison signal presented first. The legend 'both' refers to the mean proportion over both orders.

Experiment 15a) has 20 trials per cell for each standard fundamental frequency and presentation order combination, and 60 trials for the mean data. For the data pooled over the two presentation orders, there were 20 trials per cell for each fundamental frequency, and 120 for the mean data.

The number of trials per cell for Experiment 15b) was 40 for each standard signal, and 120 for the mean data.

Observer A.F. Experiments 15a) and 15b)					
		\hat{n}			
T_s msec	n	5	4 (H)	3	FA
20	4	0.3325	0.4300	0.2375	0.2850
		0.2350	0.4450	0.3200	0.2775
		0.1875	0.9250	0.0000	0.0938
		0.0000	0.8750	0.0125	0.0063
		0.5000	0.3125	0.0000	0.2500
		0.2292	0.7042	0.0042	0.1167
25	4	0.1725	0.6900	0.1375	0.1550
		0.1875	0.6300	0.1825	0.1850
		0.2625	0.9000	0.0000	0.1313
		0.0000	0.9375	0.0000	0.0000
		0.1625	0.3875	0.0000	0.0813
		0.1417	0.7417	0.0000	0.0708
40	4	0.1375	0.8025	0.0600	0.0988
		0.0975	0.8500	0.0525	0.0750
		0.1000	0.9500	0.0000	0.0500
		0.0000	0.9500	0.0000	0.0000
		0.0500	0.9500	0.0000	0.0250
		0.0000	1.0000	0.0000	0.0000
		0.0000	0.9500	0.0000	0.0000
		0.0000	0.9750	0.0000	0.0000
		0.0500	1.0000	0.0000	0.0250
		0.2000	1.0000	0.0500	0.1250
		0.1250	1.0000	0.0250	0.0750
		0.0500	0.9833	0.0000	0.0250
		0.0667	0.9667	0.0167	0.0417
		0.0583	0.9750	0.0083	0.0333
		0.1375	0.9875	0.0000	0.0688
		0.0125	0.9875	0.0000	0.0063
		0.2750	0.2875	0.0000	0.1375
		0.1417	0.7542	0.0000	0.0708

Predicted
Goldstein
Wightman
Observed (exp 15b)
 $f_o = 205$
170
258
mean
Predicted
Goldstein
Wightman
Observed (exp 15b)
 $f_o = 205$
170
258
mean
Predicted
Goldstein
Wightman
Observed (exp 15a)
 $f_o = 205$ a
205 b
205 both
170 a
170 b
170 both
258 a
258 b
258 both
mean a
mean b
mean both
Observed (exp 15b)
 $f_o = 205$
170
258
mean

Experiment 15 - Observer A.F. - Continued.

60	4	λ : 3	4 (H)	S	FA	δ'	Predicted
		0.0725	0.9175	0.0100	0.0413	4.95	Goldstein
		0.0675	0.9050	0.0275	0.0475	4.72	Wightman
							Observed (exp 15b)
		0.0500	1.0000	0.0000	0.0250	>6.30	$f_o = 205$ a
		0.0500	1.0000	0.2000	0.1250	>5.24	205 b
		0.0500	1.0000	0.1000	0.0750	>5.63	205 both
		0.0000	0.9500	0.0000	0.0000	>6.06	170 a
		0.0000	0.9500	0.0000	0.0000	>6.06	170 b
		0.0000	0.9500	0.0000	0.0000	>6.06	170 both
		0.0500	0.9500	0.0000	0.0250	5.43	258 a
		0.2000	0.8500	0.0000	0.1000	3.85	258 b
		0.1250	0.9000	0.0000	0.0625	4.52	258 both
		0.0333	0.9667	0.0000	0.0167	5.97	mean a
		0.0833	0.9333	0.0667	0.0750	4.55	mean b
		0.0583	0.9500	0.0333	0.0459	5.10	mean both
							Predicted
		0.0225	0.9775	0.0000	0.0113	6.58	Goldstein
		0.0125	0.9850	0.0025	0.0038	6.93	Wightman
80	4						Observed (exp 15b)
		0.0500	0.9000	0.0000	0.0250	4.99	$f_o = 205$ a
		0.1500	0.9500	0.1000	0.1250	4.36	205 b
		0.1000	0.9250	0.0500	0.0750	4.55	205 both
		0.0000	0.9500	0.0000	0.0000	>6.06	170 a
		0.0000	1.0000	0.0000	0.0000	>6.93	170 b
		0.0000	0.9750	0.0000	0.0000	>6.58	170 both
		0.0000	0.9500	0.0000	0.0000	>6.06	258 a
		0.2000	0.8500	0.0000	0.1000	3.85	258 b
		0.1000	0.9000	0.0000	0.0500	4.65	258 both
		0.0167	0.9333	0.0000	0.0083	5.85	mean a
		0.1167	0.9333	0.0333	0.0750	4.55	mean b
		0.0667	0.9333	0.0167	0.0417	5.04	mean both
							Observed (exp 15b)
		0.1625	0.9250	0.0000	0.0813	4.55	$f_o = 205$
		0.0000	1.0000	0.0000	0.0000	>6.93	170
		0.0375	0.7125	0.0000	0.0188	4.40	258
		0.0667	0.8792	0.0000	0.0333	4.86	mean
100	4						Predicted
		0.0525	0.9425	0.0050	0.0289	6.58	Goldstein
		0.0675	0.8850	0.0475	0.0575	4.46	Wightman
							Observed (exp 15b)
		0.2500	0.8000	0.0000	0.1250	3.40	$f_o = 205$ a
		0.2000	0.7500	0.0000	0.1000	3.44	205 b
		0.2250	0.7750	0.0000	0.1125	3.47	205 both
		0.0000	0.9000	0.0000	0.0000	>5.61	170 a
		0.0000	1.0000	0.0000	0.0000	>6.93	170 b
		0.0000	0.9500	0.0000	0.0000	>6.06	170 both
		0.0500	1.0000	0.0000	0.0250	>6.30	258 a
		0.2500	0.6500	0.0500	0.1500	2.78	258 b
		0.1500	0.8250	0.0250	0.0875	3.84	258 both
		0.1000	0.9000	0.0000	0.0500	4.65	mean a
		0.1500	0.8000	0.0167	0.0833	3.80	mean b
		0.1250	0.8500	0.0083	0.0667	4.12	mean both

Experiment 15 - Predictions and Results.
Observer A.F.

		\hat{n}			FA	d'	
		6	5 (H)	4			
T _s msec	20	5					Predicted
			0.3625	0.3325	0.3050	0.3338	0.00 Goldstein
			0.2550	0.3525	0.3925	0.3238	0.62 Wightman
							Observed (exp 15b)
			0.2625	0.9375	0.0000	0.1313	4.25 f _o =205
			0.0000	0.7125	0.1625	0.0813	3.48 170
	25	5	0.0000	0.1875	0.0000	0.0000	>3.46 258
			0.0875	0.6125	0.0542	0.0708	3.30 mean
							Predicted
			0.2550	0.5525	0.1925	0.2238	2.10 Goldstein
			0.2325	0.5300	0.2375	0.2350	1.94 Wightman
							Observed (exp 15b)
40	5	5	0.1125	0.9375	0.0335	0.0750	4.65 f _o =205
			0.0000	0.6500	0.3625	0.1813	2.60 170
			0.0000	0.2125	0.0000	0.0000	>3.53 258
			0.0375	0.6000	0.1333	0.0854	3.08 mean
							Predicted
			0.2050	0.6900	0.1050	0.1550	2.83 Goldstein
			0.1700	0.7050	0.1250	0.1475	3.02 Wightman
							Observed (exp 15a)
			0.0000	1.0000	0.0000	0.0000	>6.93 f _o =205 a
			0.0000	1.0000	0.0000	0.0000	>6.93 205 b
			0.0000	1.0000	0.0000	0.0000	>6.93 205 both
			0.0000	1.0000	0.0000	0.0000	>6.93 170 a
	5	5	0.0000	0.9000	0.0500	0.0250	4.99 170 b
			0.0000	0.9500	0.0250	0.0500	5.10 170 both
			0.0000	1.0000	0.0000	0.0000	>6.93 258 a
			0.1500	0.9000	0.0500	0.1000	4.14 258 b
			0.0750	0.9500	0.0250	0.0500	5.10 258 both
			0.0000	1.0000	0.0000	0.0000	>6.93 mean a
			0.0500	0.9333	0.0333	0.0417	5.04 mean b
			0.0250	0.9667	0.0167	0.0208	5.97 mean both
							Observed (exp 15b)
			0.0875	0.9750	0.1125	0.1000	5.10 f _o =205
			0.0000	0.8375	0.4250	0.2125	3.13 170
			0.0250	0.5375	0.0000	0.0125	4.33 258
			0.0375	0.7833	0.1792	0.1083	3.47 mean

Experiment 15 - Observer A.F. Continued.

60	5	\hat{a}	\hat{u}	\hat{c}	\hat{f}_A	\hat{d}'	Predicted
		0.1500	0.7900	0.0600	0.1050	3.51	Goldstein
		0.1225	0.7975	0.0800	0.1013	3.62	Wightman
							Observed (exp 15a)
		0.0000	1.0000	0.0500	0.0250	>6.30	$f_o = 205$ a
		0.0500	0.9000	0.0000	0.0250	4.99	205 b
		0.0250	0.9500	0.0250	0.0250	5.43	205 both
		0.0000	0.8000	0.0000	0.0000	>5.10	170 a
		0.0000	1.0000	0.2500	0.1250	>5.24	170 b
		0.0000	0.9000	0.1250	0.0625	4.52	170 both
		0.0000	0.9000	0.0000	0.0000	>5.61	258 a
		0.1000	0.9500	0.2000	0.1500	4.24	258 b
		0.0500	0.9250	0.1000	0.0750	4.55	258 both
		0.0000	0.9000	0.0167	0.0083	5.61	mean a
		0.0500	0.9500	0.1500	0.1000	4.58	mean b
		0.0250	0.9250	0.0833	0.0542	4.89	mean both
80	5						Predicted
		0.0375	0.9575	0.0050	0.0213	5.81	Goldstein
		0.0200	0.9625	0.0175	0.0188	5.81	Wightman
							Observed (exp 15a)
		0.0000	0.9000	0.0000	0.0000	>5.61	$f_o = 205$ a
		0.0000	0.9000	0.0500	0.0250	4.99	205 b
		0.0000	0.9000	0.0250	0.0125	5.61	205 both
		0.0000	0.8500	0.0500	0.0250	4.70	170 a
		0.0000	0.9000	0.1500	0.0750	4.31	170 b
		0.0000	0.8750	0.1000	0.0500	4.52	170 both
		0.0000	0.9000	0.0000	0.0000	>5.61	258 a
		0.1000	0.8000	0.1500	0.1250	3.40	258 b
		0.0500	0.8500	0.0750	0.0025	>5.32	258 both
		0.0000	0.8833	0.0167	0.0083	5.49	mean a
		0.0333	0.8667	0.1167	0.0750	4.13	mean b
		0.0167	0.8750	0.0667	0.0417	4.67	mean both
							Observed (exp 15b)
		0.0750	0.8750	0.1125	0.0938	5.19	$f_o = 205$
		0.0000	0.9750	0.0875	0.0438	5.77	170
		0.0000	0.8500	0.0125	0.0083	5.32	258
		0.0250	0.9333	0.0708	0.0479	4.89	mean
100	5						Predicted
		0.0525	0.9425	0.0050	0.0288	5.32	Goldstein
		0.0675	0.8850	0.0475	0.0575	4.46	Wightman
							Observed (exp 15a)
		0.1000	1.0000	0.0500	0.0750	>5.63	$f_o = 205$ a
		0.1000	0.8000	0.0000	0.0500	4.14	205 b
		0.1000	0.9000	0.0250	0.0625	4.52	205 both
		0.0000	0.8500	0.1500	0.0750	4.02	170 a
		0.0000	0.8500	0.2500	0.1250	3.63	170 b
		0.0000	0.8500	0.2000	0.1000	3.85	170 both
		0.0000	1.0000	0.0000	0.0000	>6.93	258 a
		0.2000	0.8000	0.1000	0.1500	3.28	258 b
		0.1000	0.9000	0.0500	0.0750	4.31	258 both
		0.0333	0.9500	0.0667	0.0500	5.10	mean a
		0.1000	0.8167	0.1167	0.1083	3.63	mean b
		0.0667	0.8833	0.0917	0.0792	4.19	mean both

Experiment 15 - Predictions and Results.
Observer G.S.

		\hat{n}				
T, msec	n	5	4 (H)	3	FA	d'
20	4					Predicted
		0.2650	0.5250	0.2100	0.2375	1.94 Goldstein
		0.2075	0.5350	0.2575	0.2325	2.02 Wightman
						Observed (exp 15b)
		0.0000	0.1375	0.3875	0.1937	-1.12 $f_o = 205$
		0.0500	0.2750	0.3500	0.2000	1.19 170
		0.1375	0.4125	0.5125	0.3250	0.98 258
		0.0625	0.2750	0.4167	0.2396	0.80 mean
						Predicted
		0.1900	0.6975	0.1125	0.1513	2.93 Goldstein
		0.1650	0.6800	0.1550	0.1600	2.80 Wightman
						Observed (exp 15b)
25	4					
		0.0125	0.2875	0.3500	0.1813	1.41 $f_o = 205$
		0.1000	0.2250	0.4125	0.2563	-0.71 170
		0.1375	0.4750	0.4625	0.3000	1.47 258
		0.0033	0.3292	0.4083	0.2458	1.09 mean
						Predicted
		0.0850	0.9025	0.0125	0.0488	4.65 Goldstein
		0.0550	0.9175	0.0275	0.0413	4.95 Wightman
						Observed (exp 15a)
		0.1500	0.9500	0.2000	0.1750	4.07 $f_o = 205$ a
		0.3000	0.9500	0.0000	0.1500	4.24 205 b
		0.2250	0.9500	0.1000	0.1625	4.18 205 both
40	4	0.0000	0.8500	0.1000	0.0500	4.36 170 a
		0.1500	0.8000	0.0000	0.0750	3.80 170 b
		0.0750	0.8250	0.0500	0.0625	4.14 170 both
		0.2000	0.3500	0.2000	0.2000	1.55 258 a
		0.0500	0.4000	0.0500	0.0500	3.02 258 b
		0.1250	0.3750	0.1250	0.1250	2.16 258 both
		0.1167	0.7167	0.1667	0.1417	3.05 mean a
		0.1667	0.7167	0.0167	0.0917	3.42 mean b
		0.1417	0.7167	0.0917	0.1167	3.19 mean both
						Observed (exp 15b)
		0.0250	0.5000	0.3500	0.1875	2.12 $f_o = 205$
		0.0000	0.4500	0.3125	0.1563	2.16 170
		0.1250	0.3500	0.5750	0.3500	0.00 258
		0.0500	0.4333	0.4125	0.2313	1.67 mean

Experiment 15 - Observer G.S.

60	4	$\hat{\lambda} = 3$	$k (H)$	\leq	f_A	d'	Predicted
80	4	0.0475	0.9450	0.0075	0.0275	5.43	Goldstein
		0.0550	0.9250	0.0200	0.0375	5.04	Wightman
							Observed (exp 15a)
		0.0000	0.9500	0.2500	0.1250	4.36	$f_o = 205$ a
		0.1500	1.0000	0.0000	0.0750	>5.63	205 b
		0.0750	0.9750	0.1250	0.1000	5.10	205 both
		0.0000	0.4000	0.1000	0.0500	3.02	170 a
		0.1000	0.8000	0.0000	0.0500	4.14	170 b
		0.0500	0.6000	0.0500	0.0500	3.51	170 both
		0.0000	0.2500	0.2000	0.1000	1.95	258 a
		0.2000	0.3500	0.1000	0.1500	1.91	258 b
		0.1000	0.3000	0.1500	0.1250	1.88	258 both
		0.0000	0.5333	0.1833	0.0917	2.89	mean a
		0.1500	0.7167	0.0333	0.0917	3.42	mean b
		0.0750	0.6250	0.1083	0.0917	3.16	mean both
							Predicted
		0.0300	0.9700	0.0000	0.0150	5.97	Goldstein
		0.0250	0.9625	0.0125	0.0188	5.81	Wightman
							Observed (exp 15a)
		0.0000	0.9000	0.0500	0.0250	4.99	$f_o = 205$ a
		0.1000	0.9500	0.0000	0.0500	5.10	205 b
		0.0500	0.9250	0.0250	0.0375	5.04	205 both
		0.0000	0.8000	0.2500	0.1250	3.40	170 a
		0.1500	0.7000	0.0000	0.0750	3.45	170 b
		0.0750	0.7500	0.1250	0.1000	3.44	170 both
100	4	0.3000	0.2000	0.1500	0.2250	-0.76	258 a
		0.2000	0.4000	0.1500	0.1750	1.87	258 b
		0.2500	0.3000	0.1500	0.2000	1.31	258 both
		0.1000	0.6333	0.1500	0.1250	2.85	mean a
		0.1500	0.6833	0.0500	0.1000	3.21	mean b
		0.1250	0.6583	0.1000	0.1125	3.08	mean both
							Observed (exp 15b)
		0.0000	0.8125	0.1250	0.0625	4.05	$f_o = 205$
		0.0250	0.5125	0.0625	0.0438	3.44	170
		0.0750	0.3125	0.2750	0.1750	1.51	258
		0.0333	0.5458	0.1542	0.0938	2.94	mean
							Predicted
		0.0000	0.9825	0.0175	0.0088	6.58	Goldstein
		0.0325	0.9525	0.0150	0.0238	5.68	Wightman
							Observed (exp 15a)
		0.1000	0.7000	0.3500	0.2250	2.50	$f_o = 205$ a
		0.2000	0.9000	0.0500	0.1250	3.92	205 b
		0.1500	0.8000	0.2000	0.1750	3.11	205 both
		0.0500	0.6000	0.3500	0.2000	2.35	170 a
		0.1500	0.8500	0.0500	0.1000	3.85	170 b
		0.1000	0.7250	0.2000	0.1500	3.02	170 both
		0.1500	0.5000	0.4500	0.3000	1.54	258 a
		0.2000	0.4000	0.2000	0.2000	1.74	258 b
		0.1750	0.4500	0.3250	0.2500	1.63	258 both
		0.1000	0.6000	0.3833	0.2417	2.15	mean a
		0.1833	0.7167	0.1000	0.1417	3.05	mean b
		0.1417	0.6583	0.2417	0.1917	2.58	mean both

Experiment 15 - Predictions and Results.
Observer G.S.

		\hat{n}			FA	d'	
T_s msec	n	6	5 (H)	4			
20	5	0.4100	0.3025	0.2875	0.3488	-0.81	Predicted Goldstein
		0.2525	0.3400	0.4075	0.3300	0.36	Wightman
							Observed (exp 15b)
		0.1500	0.4875	0.6250	0.3875	1.03	$f_o = 205$
		0.0500	0.1000	0.3125	0.1813	-1.55	170
		0.4125	0.4625	0.7000	0.5563	-0.98	258
	5	0.2042	0.3500	0.5458	0.3750	-0.60	mean
							Predicted Goldstein
		0.2375	0.5725	0.1900	0.2138	2.21	Wightman
		0.2225	0.5425	0.2350	0.2288	2.02	Observed (exp 15b)
		0.1375	0.3500	0.6375	0.3875	-0.69	$f_o = 205$
		0.1250	0.0875	0.4250	0.2750	-2.18	170
25	5	0.2125	0.4875	0.7000	0.4563	-0.54	258
		0.1583	0.3083	0.5875	0.3729	-0.88	mean
							Predicted Goldstein
		0.2200	0.6375	0.1425	0.1813	2.57	Wightman
		0.1750	0.6800	0.1450	0.1600	2.80	Observed (exp 15a)
		0.2000	0.8000	0.5000	0.3500	2.35	$f_o = 205$ a
	5	0.2500	0.8000	0.3500	0.3000	2.55	205 b
		0.2250	0.8000	0.4250	0.3250	2.43	205 both
		0.0000	0.1500	0.5500	0.2750	-1.61	170 a
		0.0500	0.4000	0.5500	0.3000	1.12	170 b
		0.0250	0.2750	0.5500	0.2875	-0.38	170 both
		0.0000	0.1000	0.2000	0.1000	0.00	258 a
40	5	0.0500	0.1500	0.1500	0.1000	1.28	258 b
		0.0250	0.1250	0.1750	0.1000	1.03	258 both
		0.0667	0.3500	0.4167	0.2417	1.27	mean a
		0.1167	0.4500	0.3500	0.2333	1.74	mean b
		0.0917	0.4000	0.3833	0.2375	1.50	mean both
							Observed (exp 15b)
	5	0.0875	0.5250	0.4500	0.2688	1.79	$f_o = 205$
		0.0250	0.0500	0.4000	0.2125	-2.42	170
		0.2250	0.4125	0.5250	0.3750	0.58	258
		0.1125	0.3292	0.4583	0.2854	0.74	mean

Experiment 15 - Observer G.S.

60	5	$\hat{\lambda}$	μ	S (%)	G	FA	d'	Predicted
		0.1750	0.7475	0.0775	0.1263	3.22	Goldstein	
		0.1600	0.7175	0.1225	0.1413	3.05	Wightman	
							Observed (exp 15a)	
		0.1000	0.7000	0.7000	0.4000	1.79	$f_o = 205$ a	
		0.2500	0.9500	0.3500	0.3000	3.51	205 b	
		0.1750	0.8250	0.5250	0.3500	2.48	205 both	
		0.0000	0.4000	0.7000	0.3500	0.77	170 a	
		0.0000	0.4000	0.3000	0.1500	2.07	170 b	
		0.0000	0.4000	0.5000	0.2500	1.44	170 both	
		0.0500	0.0500	0.2000	0.1250	-1.96	258 a	
		0.0500	0.0000	0.1500	0.1000	-3.04	258 b	
		0.0500	0.0250	0.1750	0.1125	-2.29	258 both	
		0.0500	0.3833	0.5333	0.2917	1.08	mean a	
		0.1000	0.4500	0.2667	0.1833	2.03	mean b	
		0.0750	0.4167	0.4000	0.2375	1.57	mean both	
80	5						Predicted	
		0.1450	0.7875	0.0675	0.1063	3.51	Goldstein	
		0.1250	0.7725	0.1025	0.1138	3.43	Wightman	
							Observed (exp 15a)	
		0.2000	0.8500	0.2000	0.2000	3.23	$f_o = 205$ a	
		0.2000	0.7500	0.4000	0.3000	2.36	205 b	
		0.2000	0.8000	0.3000	0.2500	2.76	205 both	
		0.0000	0.3000	0.6000	0.3000	0.00	170 a	
		0.1000	0.3500	0.4000	0.2500	1.20	170 b	
		0.0500	0.3250	0.5000	0.2750	0.84	170 both	
		0.1000	0.2000	0.2000	0.1500	1.08	258 a	
		0.0500	0.2500	0.1500	0.1000	1.95	258 b	
		0.0750	0.2250	0.1750	0.1250	1.53	258 both	
		0.1000	0.4500	0.3333	0.2167	1.80	mean a	
		0.1167	0.4500	0.3167	0.2167	1.80	mean b	
100	5	0.1083	0.4500	0.3250	0.2167	1.80	mean both	
							Observed (exp 15b)	
		0.1000	0.6750	0.5000	0.3000	2.12	$f_o = 205$	
		0.0250	0.2125	0.3125	0.1688	0.93	170	
		0.2250	0.3875	0.6125	0.4188	-0.57	258	
		0.1167	0.4250	0.4750	0.2958	1.27	mean	
							Predicted	
		0.1350	0.8250	0.0400	0.0875	3.84	Goldstein	
		0.0825	0.8675	0.0500	0.0663	4.23	Wightman	
							Observed (exp 15a)	
		0.0000	0.6500	0.5500	0.2750	2.11	$f_o = 205$ a	
		0.3000	0.6000	0.4000	0.3500	1.65	205 b	
		0.1500	0.6250	0.4750	0.3125	1.92	205 both	
		0.0000	0.2500	0.5000	0.2500	0.00	170 a	
		0.1000	0.3000	0.5500	0.3250	-0.64	170 b	
		0.0500	0.2750	0.5250	0.2875	-0.38	170 both	
		0.0000	0.2000	0.5500	0.2750	-1.19	258 a	
		0.0000	0.1000	0.2000	0.1000	0.00	258 b	
		0.0000	0.1500	0.3750	0.1875	-0.98	258 both	
		0.0000	0.3667	0.5333	0.2667	1.17	mean a	
		0.1333	0.3333	0.3833	0.2583	1.01	mean b	
		0.0667	0.3500	0.4583	0.2625	1.13	mean both	

Methodological Terms.

- 2AFC** Two-Alternative-Forced-Choice. A discrimination method.
An observer is presented with two stimuli in two temporal intervals, and must order the two stimuli with respect to some physical or subjective difference.
- 2IAX** Two-Interval-Same-Different. An identification method,
An observer is presented with two stimuli in two temporal intervals, and must identify the two stimuli as the same or different with respect to some physical or subjective difference.
- 2IFC** Two-Interval-Forced-Choice. A detection method. An observer is presented with a stimulus in one of two temporal intervals, and must indicate the interval in which the stimulus occurred.

Physical terms.

- a amplitude.
- dB decibel.
- E The energy of a signal integrated over some time T_1 (see below).
- Δf Frequency difference - in the description of Experiments 1 to 13, Δf refers to a fundamental frequency difference for a harmonic signal.
- f Frequency in Hz
- f_c Cut-off frequency for low- or high-pass filtering.
- f(carr) Carrier frequency, the centre component of an amplitude-modulated signal.
- f_o Fundamental frequency in Hz of a harmonic complex.
- f_j Frequency in Hz of the jth component of a complex signal.
- $F|w|$ Amplitude spectrum with respect to angular frequency.
- $F|f|$ Amplitude spectrum with respect to frequency in Hz.

$F w ^2$	Power spectrum with respect to angular frequency.
$F f ^2$	Power spectrum with respect to frequency in Hz.
<u>g</u>	Modulation frequency for an amplitude- or frequency-modulated signal.
Hz	Cycles/sec.
<u>m</u>	Modulation index for amplitude-modulation. The term <u>m</u> is also used on occasion to denote a harmonic number.
<u>N₀</u>	Noise spectral-density, the noise power in a frequency band 1 Hz wide.
QFM	Quasi-frequency-modulation - a SAM signal with the phase of the centre component advanced by 90° .
SAM	Sinusoidal amplitude-modulation.
SPL	Sound pressure level - in dB relative to 0.0002 dyne/cm^2
<u>t₀</u>	The period of a harmonic signal in sec, $= 1/\underline{f_0}$ Hz.
<u>T_s</u>	Signal duration
<u>τ</u>	Delay interval in sec.
<u>θ</u>	Phase angle.
<u>w</u>	Angular frequency in radians/sec $\underline{w} = \underline{2\pi f}$.
<u> x </u>	The absolute value of <u>x</u> .

Statistical terms.

<u>d'</u>	The index of detectability in the Theory of Signal Detection.
<u>df</u>	Degrees of freedom
<u>F</u>	F ratio (also used to refer to the filter width parameter of an Energy Detection model of frequency discrimination, see below).
<u>hf</u>	Heterogeneity factor - the extent to which the variance of an observed psychometric exceeds the binomial variance. Evaluated as $\underline{\chi^2/df}$.
<u>λ</u>	The variance ratio of two series of random variables.

MSD The mean square deviation of a series of observed and predicted values.

N Number of observations.

p Probability.

pdf Probability density function.

$\Phi(y)$ The area under the Normal (Gaussian) Distribution from $-\infty$ to y.

$r(x,y)$ The product-moment correlation of x and y.

$R(\tau)$ The autocorrelation function at delay τ .

$\mu(x)$ The mean of the distribution of x.

S.E. Standard Error of estimation.

t Student's t.

$\sigma(x)$ The standard deviation of the distribution of x.

$\sigma^2(x)$ The variance of the distribution of x.

v Degrees of freedom.

x A physical stimulus.

\hat{x} An internal estimate of the stimulus x.

x' The population distribution of \hat{x} .

χ^2 Chi-square.

z The Normal Deviate.

Theoretical terms.

A The amplitude of a sinusoid in units of its threshold amplitude.

DL The frequency difference limen, that frequency difference whose sign can be correctly identified with a probability of 0.76.

E' The detected energy of a signal.

ΔE The detected energy difference for a pair of signals.

\hat{f}_j An internal estimate of the frequency of the jth component

of a complex signal.

$\underline{f'_j}$ The population distribution of \hat{f}_j .

\hat{f}_0 An internal estimate of the fundamental frequency of a complex signal.

$\underline{f'_0}$ The population distribution of \hat{f}_0 .

\underline{F} The standard deviation in Hz of a Gaussian auditory filter.
See also \underline{F} ratio above.

$\underline{H(w)}$ The transfer function of an auditory filter with respect to angular frequency.

\underline{j} Subscript referring to a component of a complex signal.

\underline{k} The number of components of a complex signal.

$\underline{\mu(x')}$ The mean of the population distribution x'

$\underline{n_j}$ Harmonic number of the \underline{j} th component of a complex signal

\hat{n}_j An internal estimate of n_j

SL Sensation Level - the amplitude of a signal in dB relative to the threshold amplitude of the signal.

\hat{t}_0 An internal estimate of the period of a complex signal.

$\underline{t'_0}$ The population distribution of \hat{t}_0 .

$\underline{T_I}$ The time-constant of integration of energy.

$\underline{\sigma(x')}$ The standard deviation of the population distribution x' .

$\hat{\sigma}(x')$ An empirical estimate of the standard deviation of the population distribution x' .

$\underline{\sigma^2(x')}$ The variance of the population distribution x' .

$\hat{\sigma}^2(x')$ An empirical estimate of the variance of the population distribution x' .

$\underline{\sigma_0}$ The relative standard deviation of an internal estimate of fundamental frequency in a discrimination task where the pitches of two signals having non-coincident components are compared.

- σ_s The relative standard deviation of an internal estimate of fundamental frequency in a discrimination task where the pitches of two signals having coincident components are compared.
- σ_p The internal contribution to σ_p .
- u The frequency uncertainty factor - the number of sensory channels contributing noise divided by the number contributing information about the signal.