DISCRIMINATION OF SOUNDS IN HEARING

Harry Levitt
1964

A thesis submitted for the degree of
Doctor of Philosophy, in the Faculty of Engineering,
University of London

Department of Electrical Engineering,
Imperial College of Science and Technology,
University of London,
1964.

ABSTRACT

This thesis describes an investigation of the relative ability of a listener to separate out one of several simultaneous sounds. Experiments were performed in which a listener was required to state the apparent lateral position of a sound in the presence or absence of a second, interfering sound. If no change in the listener's performance was observed whether one or two sounds were presented, then the desired sound was said to be totally separable from the interfering sound. If a decrement in performance was observed then the relative loss in the measured ability of the listener to lateralize the desired sound was used as a measure of the degree of separability. Headphone listening with interaural time delay as the physical correlate of lateral position was used.

Preliminary experiments showed the technique to be one of great potential accuracy, but that certain problems associated with a drift in the interaural time-delay threshold needed to be resolved. The experiments were mainly with separate, but statistically identical, random noise sources. The results were interpreted in terms of a simple crosscorrelation model.

Sequential estimation techniques were considered in order to compensate for gradual drifts in the listener's performance during a test. A recently devel-

oped strategy was found to be especially suited to this task and it was used in an investigation on the relative separability of one of two physically similar, voiced messages; this being its first known practical application. The investigation showed that the degree of separability varied in a manner similar to that for the relative articulation for the corresponding listening task. The use of a sequential strategy in placing a sound image in a desired position was also demontrated.

CONTENTS

Item		Page
ABSTRACT		2
CONTENTS		4
GLOSSARY OF SYMBOLS AND ABBREVIATIONS		8
ACKNOWLEDGEMENTS		9
1.	INTRODUCTION AND REVIEW	11
1.1	Introduction	11
1.2	A Note on Publication	14
1.3	Binaural Listening - A Review	14
1.3.1	Binaural Fusion	15
1.3.2	Partial Fusion	17
1.3.3	Headphone Listening and Sound Lateralization	19
1.3.4	Free Space Listening and Sound Localization	27
1.3.5	Relative Advantages of Headphone Listening	30
1.3.6	Other Factors	33
1.3.7	Effect of Spatial Separation on Competing Stimuli	38
1.3.8	Models of Auditory Localization	41
1.3.9	Conclusions	43
2.	ON A MEASURE OF SEPARABILITY	46
2.1	Definitions	46
2.2	The Typical Lateralization Experiment	48
2.3	On the Separability of Sound Images	52
2.4	Philosophy underlying the Definitions	58
2.5	Summary	60

		Page
3.	EXPERIMENTAL TECHNIQUE AND EQUIPMENT	62
3.1	Details of Experimental Technique	62
3.2	The Delay Unit	67
3.3	Headphones	79
3.4	Data Recording Unit	83
3.5	Ancillary Equipment	86
3.6	Concluding Remarks	87
4.	STATISTICAL TECHNIQUES	90
4.1	Introduction	90
4.2	On the Estimation of Difference Limens	91
4.3	Sequential Estimation Techniques	105
4.4	Summary and Conclusions	113
5.	SERIES A : PRELIMINARY EXPERIMENTS	115
5.1	Introduction	115
5.2	Experiments A1 and A2	117
5.3	Experiment A3	129
5.4	Subjective Tests for Lack of Balance in the Equipment	132
5.5	Stability of Performance during a Test	135
5.6	Towards Optimum Data Allocation	138
5.7	Checking Statistical Assumptions	141
5.8	Experiment A4. Images away from the Medial Plane	143
5.9	Summary and Conclusions	150
6.	NOISE IMAGES AND THE CORRELATION PROCESS	154
6.1	Introduction	154
6.2	Experiment B1. Ability to Detect whether One or Two Sources are Presented.	155
6.3	Experiment B2. The Separability of Multiple Noise Images	159

		Page
6.4	Interpretation of Results in Terms of a Correlation Model	162
6.5	Experiment B3. Effect of Bandwidth on Ability to Discriminate between One and Two Noise Sources	168
6.6	Experiment B4. Lateralization Ability and Noise Bandwidth	175
6.7	Experiment B5	178
6.8	Differences in Performance for the Various Strategies	181
6.9	Summary and Conclusions	186
7.	USE OF A SEQUENTIAL STRATEGY IN MEASURING SEPARABILITY	188
7.1	Introduction	188
7.2	Experiment C1. Separability of Two Physically Similar Voiced Messages	189
7.3	Implications of Experiment C1.	203
7.4	Experiment C2. Placing a Sound Image in the Medial Plane using a Sequential Strategy	204
7.5	Summary and Conclusions	206
8.	SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH	208
8.1	Summary	208
8.2	Suggestions for Further Research	211
8.3	Conclusions	219
APPENDICES		
Appendix A1.	The Correlation Model and Low-Pass Bandlimited Noise	220

Page

Appendix A2. A Simple Method of Programming a Digital Computer for Analysis-of-Variance Computations 226

LIST OF REFERENCES 242

GLOSSARY OF COMMONLY USED SYMBOLS AND ABBREVIATIONS

β = average slope of true lateralization curve = $1/\sigma$

μ = midpoint of true lateralization curve

σ = standard deviation of derived probability density function ($= 1/\beta$)

σ_{ε}^{2} = error variance

δI = interaural intensity difference

δT = interaural time difference

τ = relative time displacement (as occurs in cross- and autocorrelation)

ω = angular frequency

b = estimate of β

c.p.s. = cycles per second

DL = difference limen

F_c = cut-off frequency, cycles per second

k_i = various constants

m = estimate of μ

n_i = number of presentations at level x_i

N = number of consecutive identical responses before a change in x_i (Wetherill strategy)

p = proportion

s = estimate of σ

T = averaging time

$V_{(---)}$ = variance of (---)

w = probit weight

W = bandwidth (in radians)

x_i = dosage level (i.e. value of δT used in a given presentation)

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to his Supervisor, Professor E.C. Cherry, for his guidance and encouragement during the course of the work. Thanks are also due to Dr. G.M.Jenkins of the Royal College of Science and to Dr. G.B.Wetherill of Birkbeck College for several helpful discussions on statistical matters; to Edith L.R.Corliss of the National Bureau of Standards, Washington, D.C., for much helpful advice on methods of measurement, and to the author's numerous colleages at the Imperial College, particularly Dr. H.B.Voelcker,Jr., Dr.A.Sekey and Mr. R.E.C.White, for many informative and critical discussions.

The skilled technical assistance of Mr. A. White and others on the technical staff of the Imperial College is gratefully acknowledged. Special acknowledgement is due to Mr. B. Cutler for his help in the construction of the data-recording unit. The librarians at the Imperial College are to be thanked for their efficiency in tracing obscure publications; the secretarial staff for efficient and capable servies, and Miss M. Hudgell and Miss J. Adams, in particular, for their speed and accuracy in typing this thesis.

The author is heavily indebted to the many friends and colleagues who gave up so much of their free time in order to participate in the experiments. Their assistance was invaluable.

This project would not have been possible without the generous financial assistance made available by the Ministry of Aviation, the Philips Electrical Company, Eindhoven and the University Grants Committee. The writer is especially grateful to the Trustees of the Beit Scientific Research Fellowships for the award of a fellowship which supported him in the early stages of the work, and to the Medical Research Council for subsequent financial support. The University of London Computer Unit are to be thanked for the use of their facilities.

CHAPTER 1 INTRODUCTION AND REVIEW

1.1 Introduction

In July 1960 Cherry and Bowles published a note in the Journal of the Acoustical Society of America suggesting a method for measuring the ability of a listener to separate out one of two simultaneous sounds. The investigation described in this thesis is a more detailed study of the proposed method, its implications and practical feasibility.

The technique is based on the following argument: consider two sound sources each of which, when presented alone through a pair of headphones results in the listener hearing a single sound image occupying a distinct region in subjective space. If the two sources are presented simultaneously, then depending on whether the sounds are "separable" or not, the listener will hear two images or one image respectively. If the two sounds are totally separable (to a human listener), then each sound image will occupy its own characteristic region in subjective space whether the sounds are presented simultaneously or one at a time. The spatial position of a sound image can be measured comparatively easily and with considerable accuracy; this technique therefore provides a convenient, albeit indirect, method of determining whether a listener can separate out one of several sounds.

Central to this argument is the belief that any process which involves the *spatial* separation of sound images must logically include the separation of these images *as images*, that is; ".... one image cannot be subjectively placed in a distinct position from another, unless these images have first been separated (as gestalten)." (Cherry and Bowles, 1960). Note that the converse is not necessarily true; image separation does not always involve spatial separation, as is clear in the case of two voices emanating from a single loudspeaker. Furthermore, it should be noted that the situation can arise where the two sound sources produce images which are spatially apart, but whose relative positions are different for single and joint presentations. In this case it is argued that the stimuli produce images which are separable, but only to a certain degree. A measure for the "degree of separability" is postulated which is based on the relative interference of one image on the spatial resolution and/or position of the other.

The above argument has, by way of introduction, been framed in general terms. In the following chapter the concepts are given exact expression. Nevertheless, even in general terms it is apparent that the above interpretation of separability is only one of many. Moreocver, even within the different schools of thought it is accepted that there is a whole hierarchy of processes which are involved in the separation of two or more sounds. The author believes that the type of separation measured by the Cherry-Bowles technique - namely that of image separation - is one that is fundamental to the whole process of aural perception. For as Cherry has so aptly remarked:

" .... the fundamental problem of aural perception is not that of locating sounds in space nor that of recognizing voices and other images, but rather that it is the separation and creation of these images themselves from the great mixture of sounds which falls upon the two ears. The brain separates and forms these images, especially those of voices, with remarkable speed and facility and, as a consequence of this separation, the concept of space arises." (Cherry, 1963).

Finally, it may be of interest to consider why a project of this nature should be undertaken in a department of electrical engineering. The transmission of information has long been a problem of interest to engineers. In the early days research centred on the then formidable problem of developing the means of communication. Today this emphasis on hardware alone is receding and engineers are taking a much broader view of communication problems. Furthermore, recent advances in mathematics and in science generally, (e.g. Wiener 1948, Shannon and Weaver 1949) have led to a partial breakdown in the barriers dividing the various disciplines. For example, the concepts of Information Theory have proved to be of great value to psychologists (Broadbent 1958, Quastler 1955), while at the same time engineers are finding ever increasing application of discoveries made in other fields (e.g. Rosenblatt 1962). It is to be hoped that this trend will continue.

1.2 A Note on Publication

It is intended to publish the results of Chapters 5, 6 and 7 as a logical extension to the work of Cherry and Bowles (1960). A second paper on the use of a sequential strategy for measuring difference limens (Chapter 4) is also in preparation. The essence of an ancillary paper on a method of programming digital computers for analysis-of-variance computations is given in Appendix A2. Some work relevant to Chapter 3, (on the analysis of repetitive, linear networks) has recently been published, (Levitt 1961, 1962).

1.3 Binaural Listening - A Review

In compiling this review particular attention has been paid to those aspects of binaural listening of direct relevance to the current investigation and to the experimental technique in particular. The areas of greatest interest were:

a) Binaural Fusion

b) Sensitivity of a listener to interaural stimulus differences.

c) Differences between free-space and headphone listening.

d) Effects of spatial separation on competing stimuli.

Even within this restricted range of interest the relevant literature was found to be considerable and only the more important papers were considered. In quoting references earlier work was generally not referred to if covered by the current reference. A history of early experimental work in the field has been compiled by Boring (1942).

1.3.1 Binarual Fusion

When two identical stimuli are applied dichotically (i.e. one to each ear, by means of headphones) the listener almost invariably claims to hear a single sound. Under these circumstances it is said that the two stimuli have combined to form a single, fused "sound image" (see section 2.1). The limits of the fusion process have not been fully explored, but the following factors are known to be of importance:

a) Interaural time difference.

If an interaural time difference is introduced a single image is heard having an apparent position somewhat to one side of the head. If, however, the interaural delay exceeds a critical value the image splits up into two parts, one at each ear. For speech this critical delay has been observed to be 10-20 milliseconds (Cherry and Taylor 1954). Similar figures have been obtained by Blodgett, Wilbanks and Jeffress (1956) for low-frequency noise bands (approx. 100 to 200 c.p.s.). Somewhat lower values (approx. 5 millisecs.) have been observed for low-pass, bandlimited noise of approximately 4000 c.p.s. bandwidth (Blodgett et al 1956, Foote 1963).

b) Low-frequency envelope.

The ear exhibits two distinct modes of operation for frequencies above and below about 1500 c.p.s. If two different stimuli having identical low-frequency envelopes (max. freq. $<$ 1500 c.p.s.) are applied dichotically a single fused image is heard. Under these conditions large spectral differences between stimuli can be tolerated without a breakdown in fusion. This effect has been demonstrated by Leakey, Sayers and Cherry (1958) for two identically modulated sinewaves of different frequency; by Broadbent and Ladefoged (1957) for the oscillatory responses of two different resonant circuits pulsed in synchronism; by David and Guttman (1959) for simultaneous bursts of uncorrelated noise; and by Deatherage (1961) for differentially filtered clicks (high frequencies to one ear, low frequencies to the other), provided the frequency difference was not large. It has also been shown by Broadbent (1955) that differentially filtered speech applied dichotically leads to a single sound image. The latter effect is presumably due to the fact that the temporal structure of the speech waveform is common to both

c) Switching stimuli between the ears.

Cherry (1953) showed that if a speech message is switched rapidly between the ears (switching period less than 0.1 seconds) a single, fused image is heard. A slow switching rate gives rise to a lateral image which alternates from side to side. The above finding has been confirmed by Schubert and Parker (1955) and by Bocca (1961) who has advocated various clinical applications of the technique .

1.3.2 Partial Fusion

If two similar but not identical stimuli are presented dichotically then in some cases partial fusion may occur. For example, if the stimuli are two partially correlated random noises then a diffuse image rather than a distinct, centrally placed image is heard. Jeffress and Blodgett (1962a) have shown that listeners can provide better than chance lateralization judgements for two noise stimuli having an index of correlation as low as 0.1. Pollack and Trittipoe (1959a, 1959b) have measured just noticeable differences (j.n.d.) in degree of correlation for various reference correlations. There was a marked increase in sensitivity as the degree of correlation increased; the extremes of the range were j.n.d. = 0.4 for a reference correlation of zero and a j.n.d. = 0.04 for a reference correlation of unity. Factors which affected the listener's ability in this task were sound level, frequency spectrum, stimulus duration and interaural balance. For very short bursts of noise, overall stimulus duration is the prime determinant of performance (Pollack 1960).

Another form of partial fusion has been demonstrated by Deatherage (1961) who showed that if a pair of clicks were filtered, high-pass to one ear and low-pass to the other, where the frequency difference was large then two separate click-images were heard. Each of these images could be brought to the medial plane by appropriate adjustment of the interaural delay between the original clicks. Guttman (1962) has also shown that more than one click-image can be lateralized if the interaural time delay exceeds 3 - 5 millisecs. Multiple click images were similarly obtained in a study of monaural masking using a 3-click paradigm (Guttman, van Bergeijk and David 1960, see also section 1.3.6.1).

One of the earliest observations of multiple images was that of a pure tone subjected to an interaural intensity difference (see Banister 1927 p.437), but it is only recently that this effect has been measured with any precision (Whitworth and Jeffress 1961).

An interesting form of partial fusion is that in which two distinctly different stimuli are presented dichotically and where there is a tendency for one of the sounds to be localized towards the centre of the head. The effect has been described by Teas (1962) for clicks and noise (or pure tones) as the two dichotic stimuli and by Egan (1948) for speech and noise. A similar effect has been observed by Thurlow and Elfner (1959) for two harmonically related tones presented dichotically and by Thurlow and Elfner (1961) for continuous random noise in one ear and pulsed random noise (same source) in the other.

1.3.3 Headphone Listening and Sound Lateralization

With headphone listening the binaural sound image (assuming fusion occurs) appears to come from within, or near to, the head. The most prominent spatial co-ordinate of the image is its lateral position - hence the common practice of using lateralization rather than localization judgements under these conditions. It has been pointed out by Jeffress and Taylor (1961), however, that repeatable judgements of apparent azimuth position are possible with headphone listening.

1.3.3.1 Properties of the Binaural Sound Image

Generally it would appear that two ears are better than one. The auditory threshold for the binaural listening situation has been found to be 3 to 6 dB lower than that for the monaural case (Hirsh 1948), although it is possible that the observed effect may be partly due to a statistical bias in the measurements (Smith and Licklider 1949). The loudness of a binaural sound image is also greater than that for monaural stimulation and the differene in loudness appears to increase with stimulus level (Reynolds and Stevens 1960). Pitch discrimination is apparently improved, although this effect may be a result of the increased loudness of the binaural sound image (Pikler and Harris 1955).

1.3.3.2 Factors influencing Lateral Position.

The apparent position of the binaural sound image is primarily dependent on interaural time and/or intensity differences. Interaural intensity difference is the dominant cue at high frequencies ($>$ 1500 c.p.s.) whereas at low frequencies both time and intensity differences are of importance, the latter becoming progressively less so at lower frequencies, (e.g. see

Moushegian and Jeffress 1959, Mills 1960, David and Guttman 1959 and references cited therein). To a certain extent it is possible to compensate for the effect of interaural delay by means of an intensity difference and vice-versa. The nature of this trading relation, however, is highly complex and depends on the frequency spectrum, sensation level and the magnitude of the intensity-time differences, (e.g. see Cherry and Sayers 1959, David, Guttman and van Bergeijk 1958, Whitworth and Jeffress 1961, Harris 1960, Deatherage and Hirsh 1959 and references cited therein). There appear to be no simple rules, empirical or otherwise, for predicting this relationship. Furthermore, even if it is possible to balance interaural time and intensity differences, the interaction of these parameters may influence other properties of the sound image. Sayers (1957), for example, has shown that there is a marked change in a listener's performance if both interaural time and intensity differences are used.

The auditory system is particularly sensitive to transients and when a binaural signal is presented which contains conflicting lateralization information then the cues contained in the initial or transient portion of the waveform will dominate. It has been estimated that the effect of a brief transient cue may persist for several hundred milliseconds after its occurrence

(Franssen 1960, Tobias and Schubert 1959).

1.3.3.3 Precision of Centering a Sound Image

A matter of considerable interest in this study is the precision with which a listener can lateralize a sound image about his medial plane. A variety of techniques have been employed in measuring interaural time and intensity thresholds (Zwislocki and Feldman 1956). The methods which have yielded the smallest difference limens in either interaural time or intensity difference appeared to be those employing either a rapid reversing of channels (Jeffress and Blodgett 1962b, Kikuchi 1957) or those obtained by plotting the psychometric function from a large number of forced-choice, binary judgements (Klumpp and Eaddy 1956, Bowles 1960). Under optimum conditions and using impulsive stimuli (e.g. clicks, noise) difference limens of 10 -20 microseconds have been obtained; wide variations between subjects have also been observed. The difference limens for pure tones were somewhat higher being of the order of 30 - 50 micro seconds (Zwislocki and Feldman 1956). The least accurate method of measurement appears to have been that in which the stimulus was presented continuously and the subject adjusted the image position of his own accord. In this case the standard ~~error~~ deviation of successive adjustments was found to be of the order of 50 microseconds for impulsive stimuli (Jeffress and Blodgett 1962b, Kikuchi 1957).

Although the auditory system is very sensitive to clicks and other impulsive stimuli, overall duration of the stimulus is an important factor. Tobias and Zerlin (1959) have shown that the difference limen in interaural delay for noise bursts decreases rapidly with duration, reaching an asymptote of approximately 10 microseconds at a burst duration of a second or more.

Measurements of the difference limens in interaural intensity have also been carried out but not to the same extent. Mills (1960) has estimated the difference limen for tone bursts ($1\frac{1}{4}$ seconds duration, 20 milliseconds rise time) to be about 0.5 to 1.0 dB, depending on the frequency. Stimuli at 50dB above Minimum Audible Pressure were used.

The remarkable sensitivity of the auditory system to minor differences between dichotic stimuli was recognized at an early date, although the equipment used at the time was comparatively crude by present day standards. Even as recently as 1930 brass tubing was being used in order to achieve interaural delays (Bekesey 1930 , see also Jongkees and van der Veer 1958). With the advent of modern electro-acoustic devices great precision has been achieved in specifying and controlling the auditory stimuli. The results quoted in this section have all been of recent origin. References to earlier investigations along these lines may be found in Boring (1942), Kitagawa and Shintaku (1957) and Rosenzweig (1961).

1.3.3.4 Stability of Image Position

Apart from interaural time and intensity differences, which are the prime determinants of lateral position, other factors may influence the apparent position of the sound image:

a) Muscular activity.

Clenching the jaw on one side or turning the head sideways tends to pull the image towards the opposite side (Bowles 1960).

b) Vestibular Function.

It has been demonstrated that rotation or acceleration of the body immediately before or during a test can lead to angular displacements of apparent image position of as much as 20° (Arnhoult 1952, Clark and Graybiel 1949, Graybiel and Niven 1951). It should be noted that the above experiments and several of those described below were carried out under free-space conditions and not with headphone listening.

c) Interaction with other modalities.

Interaction between auditory and visual cues have been observed on numerous occasions (Arnhoult 1952 and references cited therein). The influence of visual cues on auditory localization, however, has received little attention (Jackson 1953, Arnhoult 1952). In the few experiments which have been carried out along these lines it was found that visual cues had a marked effect; in certain situations involving conflicting auditory and visual information the visual cues were sometimes found to dominate. (Jackson 1953, Witkin, Wapner and Leventhal 1952).

d) Asymmetric stimulation.

The application of a monaural stimulus in addition to a binaural stimulus can cause a marked shift in the apparent position of the binaural sound image. For example, if a portion of the basilar membrance in one ear is activated by noise of a given frequency band then a simultaneous click applied to both ears will only be perceived in the "noisy" ear after it has traversed the masked portion of the basilar membrane, resulting in a click-image very much to one side. Schubert and Elpern (1959) have used this technique for estimating the velocity of the travelling wave in the cochlea. The effect of monaural masking on binaural image position has also been demonstrated by Butler and Naunton (1962 - see also Schubert 1963) and by Raab and Osman (1962). In ~~all these~~ most cases a shift in the position of the binaural image towards the unmasked side was observed. This experimental artifice has also been exploited by Deatherage and Hirsh (1959) and by Flanagan, David and Watson (1962) in psychophysical studies of cochlea mechanics.

e) Prolonged stimulation.

If one ear is fatigued or impaired then identical stimulation at the two ears will not give rise to a centrally located image. Prolonged stimulation in one ear can produce this effect and it has been suggested that a lateralization technique be used to measure perstimulatory fatigue in one ear (Wright 1960).

A similar effect has been demonstrated by Krauskopf (1954) using a loudspeaker placed to one side of the subject's midline. After prolonged stimulation (several minutes) there is a shift in the listener's impression of the position of his medial plane. The shift is towards the loudspeaker and the induced distortion decays with time. Krauskopf has made the important observation that prolonged stimulation by a centrally placed loudspeaker can reduce the variability of a listener's judgements of medial plane position.

f) Spectral differences.

Although not investigated in great detail (for headphone listening) it has been shown that with monaural presentations the associated sound image can be localized to a certain extent depending on the spectral content of the stimulus (Mouzon 1955). With free-space listening spectral differences are obviously of great importance (see section 1.3.4).

g) Unusual stimulation.

By relaying to the ear the information picked up at two points in the sound field (e.g. by use of microphones) several unusual effects have been observed. If the two points are located near the two ears but subject to an angular displacement then the listener's impression of the sound field is rotated by a similar amount (Held 1955). Reversing the connections to the two ears can give rise to a front-back reversal of image position (Jongkees and van der Veer 1958).

1.3.4 Free-Space Listening and Sound Localization

With free space listening (as opposed to headphone listening) the apparent position of a sound image is not restricted to any particular region in space - hence the use of the term localization rather than lateralization. The sound image need not necessarily appear to occupy the same position as the sound source, and it is possible to create the illusion of a single image using two or more sources (the Stereophonic Effect). The apparent position of a stereophonic sound image is dependent on inter-channel time and/or intensity difference, amongst other factors (Snow 1954, Sandel, Teas, Feddersen and Jeffress 1955, Leakey 1957). As before, the one parameter can be traded for the other, to a certain extent. The magnitude of the inter-channel time or intensity disparity required to produce a just noticeable difference in image position, however, is some 4 - 5 times larger than the corresponding interaural differences.

Although free-space listening as such is not of direct concern in this investigation it is of interest to consider the essential differences between headphone and free-space listening. In the latter there are at least three important additional factors:

(a) Anatomy of head and external ear.

It is clear that the reflective properties of the head and external ear play an important role, particularly with respect to front-back localization (Norlund 1962, Jongkees and vander Veer 1958). Several investigators believe that it is this selective diffraction of sound waves, particularly at high frequencies, which permits the illusion of depth (Franssen 1960, p.56). Certainly it is true that depth perception is easier when the sound source is closer to the listener. (Edwards 1955) and where the frequency-dependent diffraction is relatively greater. It is also believed that the diffractive properties of the external ear provide sufficient information for the coarse localization of sounds using only one ear (Franssen 1960, Mouzon 1955, Mathes 1955). Studies with artificial heads have shown that the inter-aural intensity relations at the eardrum (for a single sound source in an anechoic room) is extremely complex at high frequencies (Norlund 1962).

(b) Head movements.

The head is continually moving in short angular jerks of up to 5° peak to peak (Jongkees and van der Veer 1958) and it is believed by several workers that the cues arising from these minute movements enable a listener to form an external sound image (Franssen 1960). De Boer (reported in Franssen 1960) has provided some convincing evidence in support of this hypothesis. An artificial head was used with microphones mounted at the position of the ear-drums. The output of each microphone was amplified and fed to the corresponding headphone used by the listener. If the artificial head was placed on a fixed object such as a table then the listener claimed to hear an internal sound image, but if the artificial head was rigidly mounted above the listener's head then an external sound image was perceived by the listener. A similar effect has been observed by Klensch (1948) using two funnels with rubber tubing leading the stimuli to the two ears. If large head movements were made and the funnels moved accordingly an external image was heard. If the funnels were held stationary then an internal image was heard. These experiments have been successfully repeated by Jongkess and van der Veer (1958).

Visual and vestibular cues also play a part in sound localization (Wallach 1940) although it may be that vestibular cues are comparatively unimportant unless coupled with visual cues (Jongkees and van der Veer 1958).

(c) Multiple arrivals.

Consider the stereophonic listening situation. Each loudspeaker gives rise to a signal at each ear; hence for a two-loudspeaker system there are at least two arrivals of a similar signal at each ear. In a typical room with little acoustic padding reflections from the walls can result in many more arrivals. If all the reflections occur within a few milliseconds then a single sound is heard, (see section 1.3.6.1). Ideally (for most experimental purposes) a single sound source should give rise to a single arrival at each ear. A stereophonic listening system can never provide the exact equivalent to this situation. Furthermore, whether a single source or stereophonic system is used, the experimenter has no independent control over the signals arriving at each ear.

1.3.5 Relative Advantages of Headphone Listening

The great advantage of headphone listening is that it affords the experimenter considerable control over the stimuli. Not only is greater precision possible in specifying and presenting the stimuli, but the signals to each ear can be controlled independently. This control, however, is not complete; coupling of the earphone to the listener's head presents certain problems, and some crosstalk between the ears is inevitable (Békesy 1948, Zwislocki 1953). In this respect the use of circumaural headphones is of particular value in that it improves coupling and reduces crosstalk. Furthermore, reasonably good ambient noise exclusion is achieved. Shaw (1962) has shown that the frequency response of a circumaural headphone as measured on a flat-plate coupler is within 5 - 7 dB of that obtained with a probe microphone just outside the entrance to the narrow part of the ear canal.

The major disadvantage of using headphones is the unnatural listening situation. Localization cues which may have resulted from head movements or frequency selective diffraction at the external ear are completely lost. Clearly, an intracranial sound image conveys less information than an external sound image. In addition, several secondary factors arise; the discomfort of wearing headphones may induce fatigue more easily and the pressure of the ear pads and/or the self-noise of the headphone enclosure may raise the auditory threshold.

A consistent difference of 6 dB in the auditory threshold has been observed between headphone and free-space listening(Munson and Wiener 1952, Rudmose 1962). Munson and Wiener have also observed a loudness disparity of 6 - 10 dB at normal listening levels. It may be noted that the variability of the auditory threshold appears to be about 50% greater if the threshold level is specified in terms of the impressed voltage across the headphones rather than in terms of the signal pressure in the ear canal. With free-space listening the variability is apparently not increased if pressure levels are measured a little way from the ear canal (Munson and Wiener 1950).

A useful compromise system may be possible in which the two loudspeakers are placed close to the listener's ears and some measure of shielding between the channels is achieved by means of head shadowing. David and Hanson (1962) claim a cross-talk ratio of 40 dB for an experimental arrangement of this type. Headphone listening is commonly used in psychoacoustic investigations, the assumption usually being made that those factors which are of importance when using headphones will similarly be of importance in the free-space listening situation.

1.3.6 Other Factors

1.3.6.1 The Precedence Effect

When two identical stimuli are presented dichotically a single sound image is heard. If a second pair of stimuli identical to the first save for a small difference in interaural time or intensity disparity, is presented shortly afterwards then a single sound image is heard having an apparent position associated with the first pair of stimuli (Wallach, Newman and Rosenzweig 1949). This effect, known as the Precedence Effect is of particular importance in stereophony and auditorium design. Although the existence of the effect was realised well over a century ago (see Fay 1958) it is only in recent years that psychophysical investigators have examined the phenomenon in any detail (David 1959 and references cited therein).

The conditions for precedence to take place are; (from Wallach et al 1949)

a) the second pair of dichotic stimuli should not lag the first by more than several milliseconds*,

b) the second stimulus pair should not exceed the first by more than 15 dB in level.

---

* the critical time delay can vary over a range from one or two milliseconds to over forty milliseconds, depending on the nature of the stimuli.

The later stimuli can influence the apparent position of the perceived sound image to a certain extent; if the two pairs of stimuli occur simultaneously then the resulting sound image will occupy a mean position midway between the positions associated with each stimulus pair.

The precedence effect is most easily demonstrated with impulsive stimuli (see also section 1.3.3.2 on the importance of transients in lateralization) and much of the current research in this area has been concerned with the temporal resolution of the auditory system to dichotic clicks or click pairs. A programme of some interest is that being carried out at the Bell Telephone Laboratories (e.g. see Guttman, van Bergeijk and David 1960). In these experiments two clicks, A and B, are presented to one ear and a "probe" click, C, to the other ear. Interaural delays are introduced so as to cause either A and C or B and C to fuse. If the spacing between A and B is sufficiently small then A masks B and only A and C will fuse. It has been found that if the pulse pattern is repeated periodically then the relative masking of B by A decreases with repetition rate. This effect has been attributed to facilitation in the neural system. Another interesting finding has been that of phantom click-images (Harris, Flanagan and Watson 1963, and references cited therein). The effect may be partly explained in terms of the oscillatory response of the basilar membrane to brief transients.

1.3.6.2 Interactions between the Ears.

Interactions between the ears can take a variety of forms. One of the earliest observations was that of binaural beats. If two tones differing slightly in frequency are applied dichotically a beat tone is heard (Fletcher 1953). The effect has only been observed for frequencies below about 1500 c.p.s. An important consequence of this finding is that it implies that at some stage in the auditory system a neural processing of the temporal features of the two stimuli must take place.

Another form of interaction that may occur is that of cross-conduction through the skull or other pathways. Bekesy (1948) has suggested that leakage and hence air-conduction may be the major source of cross-talk between the ears; Zwislocki (1953), however, has demonstrated that bone-conduction is more likely to be the cause. Zwislocki estimated the cross-talk ratio at about 40 dB for low frequencies, increasing slightly with frequency. Recent measurements by Fourcin (1964) using circumaural headphones of a similar type to those used in this investigation showed a cross-talk ratio of over 40 dB for all factors combined.

The auditory reflex can be a source of interaction between the ears in that stimulation at one ear activates the reflex on both sides. Recent measurements by Møller (1962) have indicated that for monaural stimulation the magnitude of the reflex is slightly greater in the ipsilateral ear; the reflex is greater still for binaural stimulation. An important characteristic of the reflex is that it involves a time delay of tens of milliseconds before coming into operation. Ward (1961) has suggested that activity of the auditory reflex be measured psychophysically in terms of the contralateral remote masking induced by an intense, high-frequency stimulus, preferably narrow-band noise.

Cross-masking effects have been observed on several occasions. The effects are similar in form to those of monaural masking but of smaller magnitude (e.g. see Ingham 1959). An interesting effect occurs when the contralateral masking stimulus is reduced in level; at very low levels summation rather than masking takes place (Egan 1948). Masking can occur centrally as well as peripherally and a recent model of auditory perception emphasizing this dichotomy has been proposed by M. Treisman (1963).

An unusual finding but one of some theoretical importance is that reported by Cramer and Huggins (1958). If a wide-band noise stimulus is applied simultaneously to the two ears, where the transmission characteristic of the channel to one ear has a sharp 360° phase shift over a narrow frequency band, then a pitch-like sensation is heard. Fourcin (1961) has demonstrated a similar effect using dichotic noise stimuli which, when crosscorrelated, give rise to periodic peaks in the correlation pattern resembling those associated with a pure tone. If either of the noise stimuli are presented monaurally no pitch is heard. Both these results emphasize the importance of temporal or phase information in the perception of pitch - further evidence in favour of a neural place theory of pitch perception. As with binaural beats these effects do not occur at frequencies above about 1500 c.p.s.

The comparative lack of interaction between the two auditory channels, and the ability of a listener to concentrate on the input to one or other channel at will (Cherry 1953) has provided psychologists with a useful experimental tool. For example, two speech messages may be applied dichotically and the listener required to shadow one of them; the efficiency with which the listener can carry out this task provides a good indication of his ability to "pay attention" to the desired message. The relative importance or the attention-catching value of competing stimuli may be assessed in this way (Moray 1960, A.M. Treisman 1961).

1.3.7 Effects of Spatial Separation on Competing Stimuli

It has been shown in numerous investigations that increasing the spatial separation between two or more sounds tends to reduce the relative interference between the sounds. The spatial separation may take the form of placing the actual sound sources in different positions, or it may be accomplished by introducing relative interaural time and/or intensity differences using a binaural or stereophonic listening technique.

A common method of measuring interference is that of threshold masking. The role of interaural phase or time* differences on masking has been mapped out by several investigators (Licklider 1948; Jeffress, Blodgett, Sandel and Wood 1956 and references cited therein). The effects were found to be quite large especially if one of the two competing stimuli were subjected to a 180° interaural phase reversal (Licklider 1948). This is an interesting result in that phenomenally the two sounds appear to occupy different regions in space, although the difference in localization is not very great. In the free-space listening situation Kock

---

* Note:- with pure tones a phase difference is equivalent to a time difference. With complex signals, however, the relationship between phase and time is not as simple; e.g. a 180° phase shift involves a different relative time shift at each frequency.

(1950) showed that directional differences between the sound sources were of great importance.

Generally it has been found that relative interaural phase differences can lead to a reduction in masking of the order of several decibels. Pure tones at frequencies above 1500 c.p.s. are unaffected by interaural phase differences and are an exception. Jeffress et al (1956) point out in their review that many conflicting observations in binaural masking have been obtained as a result of differences in experimental procedure. For example, substantial masking level differences have been observed when bursts of tone or bursts of uncorrelated noise were used; when continuous stimuli of the same type have been used the introduction of a relative interaural phase difference appeared to have little effect. Hirsh and Pollack (1948) have measured the effect of a relative interaural phase difference on loudness. An effect similar to that for threshold masking was observed for weak stimuli. The effects were not very marked, however, at reasonably high signal-to-noise ratios.

The spatial separation of sound images has also been shown to be of great importance in more complex perceptual tasks. Intelligibility, for example, is greatly improved if the interfering speech messages are kept spatially apart. The artifice of reversing the phase for one of two binaural messages appears to produce a greater improvement in intelligibility than that obtained by introducing a large interaural delay (over a millisecond) for one of the stimuli (Schubert 1956). Even greater improvements have been observed by Pollack and Pickett (1958) for the situation where the wanted message is presented binaurally and the unwanted messages or "noise" presented monaurally. The improvement, expressed as the equivalent increase in signal-to-noise ratio for 50% intelligibility, was found to be of the order of 12 dB if one unwanted speech message appeared in each ear; for a babble of several different voices at each ear the improvement was reduced to a 7 dB equivalent gain in S/N ratio. These figures compare favourably with those of Schubert who obtained an equivalent improvement of 5 - 6 dB by reversing the phase of one of the binaural stimuli.

A.M.Treisman (1961) has made the interesting observation that if one of three messages is shadowed, there is a decrement in the listener's performance if the two unwanted messages are separated spatially.

The role of spatial separation in allowing a listener to concentrate more efficiently on one of several sources has been emphasized by Cherry (1953, 1955) and others (e.g. Schubert and Schultz 1962). Broadbent

(1958) in an excellent review of the literature on competing messages has pointed out that selective listening to one of several messages may be improved if the different speech messages can be given distinctly different physical co-ordinates (e.g. spectral, spatial, temporal, etc.). In this respect differences in localization are of great importance and generally lead to an improvement in performance. If, however, the listener is required to respond to two or more messages simultaneously then spatial separation of the sound sources is not of much assistance.

1.3.8 Models of Auditory Localization

Psychophysical theories are as old as the science itself and innumerable models of the auditory process have been proposed. Models specifically related to localization or lateralization are in lesser abundance and several of the more important theories are here described.

In 1930 Bekesy proposed a simple physiological model to account for the effects of interaural time and/or intensity differences in auditory localization. He pictured a region of cells which were stimulated by signals from the two ears. Localization depended on which channel stimulated the most cells. Both intensity and prior arrival aided stimulation and hence the ear receiving an intense or early stimulus was favoured. van Bergeijk (1962) has proposed a modified version of the model and indicated a possible locus for it in the neural system.

In 1948 Jeffress suggested a simple, neural coincidence-detector which would account for the effect of an interaural time disparity; the possible effects of an interaural intensity difference in terms of the model were only hinted at (Jeffress 1948). This model was of great importance in that it paved the way for subsequent correlation models in which interaural time delay was the primary variable. Shortly afterwards Licklider (1951) proposed his duplex model for pitch perception and which referred to a possible correlation model for auditory localization.

The first detailed model of the lateralization process using a correlation technique was that proposed by Cherry and Sayers (1956 - see also Sayers and Cherry 1957). A running autocorrelation at the two ears followed by a running crosscorrelation was postulated. The running autocorrelation is equivalent to a running spectral analysis whereas the crosscorrelation enables the model to predict the apparent lateral position ("left" or "right") of the impressed sound source. A simple decision rule was proposed depending on the relative proportion of the correlation pattern to the right or left of the $\tau = 0$ axis. The effect of an interaural intensity difference was to truncate the pattern along the $\tau = 0$ axis.

David (1959) pointed out that intensity differences could be transformed into time differences within the neural system and several effects including the Precedence Effect were interpreted in terms of facilitation in the neural pathways. An interesting decision theory model has recently been proposed by Voelcker (1960); although based on assumptions quite unrelated to those of Cherry and Sayers the end product is a model involving a running spectral and crosscorrelation analysis of the two inputs.

Despite the diversity of the many models that have been proposed a central theme has appeared, namely that of crosscorrelation or coincidence detection. It should not be inferred, however, that a crosscorrelation process is necessarily carried out somewhere in the nervous system. Crosscorrelation is merely a useful mathematical tool which, to a first approximation, can be used to predict the listener's performance under given experimental conditions.

1.3.9 Conclusions

Perhaps the most striking aspect of the auditory system is its remarkable complexity. Certainly no single theory could account for the vast number of effects which have been observed. Even within the restricted compass of auditory localization the experimental findings are extensive. On the other hand, within a limited framework, extremely simple models can be used to account for a good many results - at least to a first approximation. A simple correlation model, for example, may cover most of the observed data in the lateralization of binaural images, but it would require a highly sophisticated extension of this model to account for the multiple images observed by Flanagan and others.

Of particular interest to this study is the remarkable sensitivity of the auditory system to interaural differences, despite the large number of extraneous factors which can influence the listener's binaural balance. In this respect the findings reviewed in sections 1.3.3 to 1.3.5 provide a good idea of the precision required of the experimental equipment and the factors to be avoided in the conduct of lateralization experiments. Unnecessary head turning, asymmetric stimulation and imperfectly balanced channels are to be avoided as far as possible.

This review, although not entirely comprehensive, does at least provide a reasonably broad picture of the psychophysics of binaural hearing. A noticeable aspect of the experimental activity in psychoacoustics is the considerable effort that has been paid to certain areas of research (e.g. stereophony) and the almost total neglect of other areas. Experimental work on the effect of small head movements in producing an external sound image, for example, has only been reported on by one group in recent years. Another aspect of binaural perception which may merit further investigation is that in which a localization effect is perceived using distinctly different, dichotic stimuli, (Egan 1948, Teas 1962). It may also be noted that whereas considerable attention has been paid to the interfering effect of one sound on another in terms of threshold masking, loudness masking and relative intelligibility, there has been little work, other than that of Cherry and Bowles (1960), on the loss in the spatial resolution of one sound image in the presence of another.

In compiling this review only psychoacoustic data directly relevant to binaural hearing, and localization in particular, was considered. Physiological knowledge was not reviewed nor were any of the general theories of hearing described. Excellent reviews of these topics may be found in Rosenzweig (1961), Bekesy (1956) and Licklider (1959). It is hoped that this brief review will provide the reader with an adequate background to the experimental investigation here described.

CHAPTER 2 ON A MEASURE OF SEPARABILITY

2.1 Definitions

As pointed out in the preceding review, when the two ears are stimulated identically the listener almost invariably claims to hear a single, centrally-placed sound. The sensation so created is referred to as a "sound image" and is said to occupy a distinct region in "subjective space". These concepts are formally defined as follows:

(a) Subjective Space:- The listener's personal impression of physical space as created by the sense of hearing.

This definition raises certain philosophical questions as to what is meant by "physical space" since, after all, knowledge of the "outside" world must in some way reach a subject through one or more senses. For the purpose of this study, "subjective space" will be regarded as that impression of space created by the sense of hearing as opposed to the impression created by all other senses combined.

(b) Sound Image:- The subjective entity, associated with a sound source, to which an attribute suchas position is ascribed.

A sound image occupies a distinct region in subjective space which may or may not coincide with the spatial position of the actual sound source. With headphone listening, assuming identical stimulation at each ear, a single sound image apparently within the head is heard and is referred to as an intracranial sound image.

The term "sound" is sometimes used in place of "sound image".

(c) Medial Plane:- The reference plane in subjective space which demarcates left from right.

(d) Localization:- The process whereby the listener ascribes an apparent position to a sound image in subjective space.

(e) Lateralization:- The process whereby the listener ascribes an apparent lateral position to an intracranial sound image.

Lateralization judgements are usually used with headphone listening since in this situation the sound image appears to come from within the head and the range of possible image positions is considerably reduced; some listeners have reported that the sound image appears to lie near the back of the head, almost along the line joining the two ears. In this study we shall be concerned only with the right- or left-handedness of the intracranial sound image.

2.2 The Typical Lateralization Experiment

The ability of a listener to lateralize a sound image about his medial plane is measured by a lateralization experiment. The following is a brief description of a typical experiment as used in this investigation. Details of the technique and equipment are discussed in the next chapter.

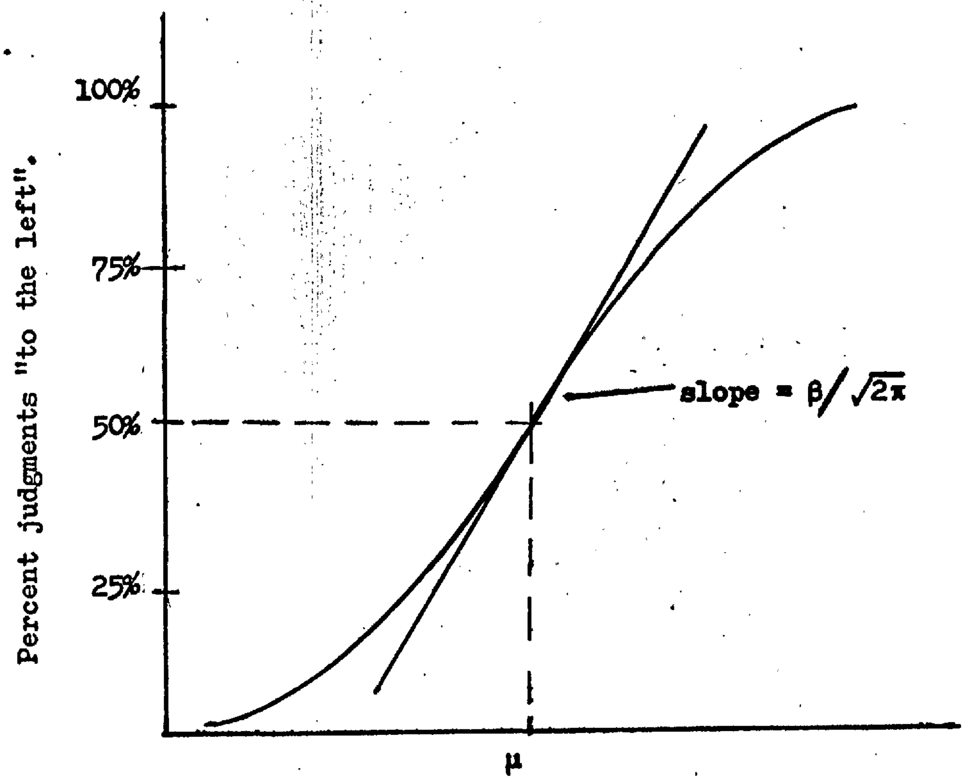
A sound is presented to a listener using a pair of matched headphones. The listener decides on the apparent lateral position of the sound image and relays his decision to the experimenter. The sound is then switched off and after a short interval a second sound is presented; the procedure is then repeated. Each cycle of events is known as a presentation. For each presentation the listener is forced to make a binary decision; either "left" or "right". No other answer is acceptable and if the listener is uncertain, he must guess.

A test consists of a large number of presentations for a given set of controlled variables, and where the physical correlate of lateral position is fixed by the experimenter immediately before each presentation. Interaural time difference (δT) is commonly used as the physical variate. The value of δT may be selected at random or in a quasi-random fashion, although recent work suggests that a sequential strategy may be more efficient, (Chapter 4).

The results of a test are plotted in the following form: proportion, or percentage of "left" decisions on the vertical axis and the corresponding value of the physical variate on the horizontal. Such a plot is known as a lateralization or judgement curve. For presentations in the vicinity of the medial plane the curve usually turns out to be sighmoidal in shape. A lateralization curve of this type is shown in Fig.2.1, (note: a negative δT indicates a delay in the channel to the left ear). The judgement curve need not necessarily be a sigmoid, other shapes have been observed under special conditions, especially if stimuli to the two ears differ markedly (Sayers and Cherry 1957, Cherry 1959).

A sigmoidal lateralization curve is interpreted in the following way; the midpoint or 50% point (i.e. μ in Fig. 2.1) corresponds to the position of the medial plane as measured in terms of the physical variate. For example, if a δT equal to μ is used the resulting sound image should appear to lie in the medial plane. The average slope of the lateralization curve is interpreted as a

FIG. 2.1 : HYPOTHETICAL LATERALIZATION CURVE



Physical correlate of lateral position (δ T).

measure of the listener's ability to lateralize the sound image about his medial plane (or, if so desired, about any fixed reference*). Clearly, the more difficult the task the greater the proportion of incorrect judgements and the shallower the lateralization curve. Average slope may be defined in a variety of ways. In this study the average slope is assumed to be a constant, $\sqrt{2\pi}$, times the slope of the curve at the mid-point. The reason for the choice of this constant is given in section 4.2.

It is important to bear in mind the essential limitations of the experimental technique. The experimenter cannot measure the subjective impressions of a listener directly but only the responses of the listener to a given set of stimuli. From observations of this type it is possible for the experimenter to make several useful inferences as to the nature of the listener's subjective impressions. It is conceivable, however, that the listener may deliberately deceive the experimenter. In this investigation it has been implicity assumed that the subjects were being honest in their judgements. It has also been assumed that both

---

* Another sound image, for example, may be used as the reference.

subject and experimenter were in agreement as to the definitions of "left" and "right".

2.3 On the Separability of Sound Images

Consider a hypothetical sigmoidal lateralization curve which has been obtained from an almost infinite number of observations, i.e. experimental errors are negligibly small. We define, (see Fig.2.1):

μ = midpoint of curve = position of the reference (e.g. the medial plane) as measured in terms of the physical correlate of lateral position.

β = ($\sqrt{2\pi}$ x slope at μ) = a measure of the listener's ability to lateralize the sound image about the given reference.

The curve fitted to a finite set of data is defined by the parameters m and b, which are estimates of μ and β respectively. Hence

m = midpoint of fitted curve = estimated position of the reference as measured in terms of the physical correlate of lateral position.

b = $\sqrt{2\pi}$ times the slope at m and is an estimate of the listener's ability to lateralize the sound image about the given reference.

It is argued that if a listener's ability to lateralize a sound image is unaffected by the presence of one or more other sounds, then the first sound is quite separable from the rest. This argument leads to the working definition of <u>separability</u>:

Consider two sound sources, A and B say, such that when either is presented alone the listener hears a sound image occupying a fixed, distinct region in subjective space.

Let β_A = expected slope of the lateralization curve for sound A about some reference (usually the medial plane), and

let $\beta_{A/B}$ = expected slope of the lateralization curve for A about the same reference but in the presence of sound B,

then if $\beta_A = \beta_{A/B}$ sound A is said to be <u>totally separable</u> from sound B. Similarly, if $\beta_B = \beta_{B/A}$ then sound B is said to be totally separable from sound A. Further, if $\beta_{A/B} \neq \beta_A$ then $(\beta_A - \beta_{A/B})$ is postulated as the <u>degree of separability</u> for sound A in the presence of B, and vice-versa for sound B in the presence of A. Note that a positive increase in $(\beta_A - \beta_{A/B})$ denotes a loss in separability.

In defining a measure for degree of separability the particular form $(\beta_A - \beta_{A/B})$ has been selected since in the practical estimation problem (using maximum likelihood estimates) a comparatively simple sampling distribution is obtained. Other monotonically related measures could be used, as for example $({}^{1}/\beta_{A/B} - {}^{1}/\beta_{A})$ or $\beta_{A/B}/\beta_{A}$ as suggested by Bowles (1960).

Note that β_A is not necessarily equal to β_B nor is $\beta_{A/B}$ necessarily equal to $\beta_{B/A}$. Hence the situation can arise where sound A is totally separable from sound B while B is barely separable from A. For example, sound A may be very loud and almost masking B entirely. In this case A is hardly affected by the presence of B, but the reverse is certainly not the case.

It should also be noted that this discussion excludes those sounds which have no fixed or distinct position in subjective space - such as a diffuse noise which appears to come from all directions, or such as a continuously moving sound image. The separability of these sounds cannot be measured directly using the proposed method, although an extension of the technique to cover such cases may be possible. Most situations of interest, however, involve sounds which satisfy the prescribed conditions.

The nature of the interaction between the spatial patterns of two simultaneous sounds can take at least two different forms; the apparent positions of the images may be altered, or their spatial resolution affected. In either case, provided the relative positions of the images for successive presentations are properly randomised during the course of a test, the affect on the slope of the lateralization curve would be the same, i.e. $\beta_{A/B}$ would be less than β_A. It is possible to determine the nature of the interference by correlating the listener's judgements with the positions of the sources δT_1 and δT_2. A fixed or constant effect would indicate a shift in image position, a random effect (i.e. an increase in error variance) would indicate a loss in resolution.

In addition to the separability of sound images, the ability of a listener to distinguish one sound from the other (i.e. known which is which) is also of interest. Another attribute, discriminability, has therefore been defined:-

If in a lateralization experiment the impressed positions of the two sound images are selected at random for each presentation and the listener can consistently select one of the images as a reference and lateralize the second image about this reference image, then the two sounds are said to be both separable and discriminable.

This study is primarily concerned with those factors which affect the separability of sound images. In certain investigations, however, the inability of the listener to identify one sound image from another complicates the experimental technique. It is for this reason that a distinction has been drawn between separability and discriminability.

Note that in the definition of discriminability the listener is not required to describe or remember after the test any of the differences between the two sound images. Discriminability is therefore fairly low down in the hierarchy of such recognition processes. Since the stimulus differential between the two ears (e.g.δT) for each source is selected at random for successive presentations, a high correlation between listener's judgements and this variate is evidence that consistent discrimination has taken place. From the nature of this correlation it is also possible to determine which sound image was used as the reference.

For sounds which are discriminable the degree of separability may be measured directly in accordance with the proposed definition. If, however, two sounds are not discriminable, but there is reason to believe that they are nevertheless separable, then certain practical difficulties arise. A case in point is that of two recordings of separate, phonetically similar passages of speech played backwards. The listener hears two separate sound images but does not know which one to lateralize since the images are not discriminable. For situations such as these the following artifice has been suggested (Bowles 1960). One of the sound images is placed in the medial plane, (the listener is allowed to adjust the associated physical variable such that, hopefully, this is the case). The test then proceeds with one image fixed in this position and the position of the other varied at random by the experimenter.

The fact that one sound image is fixed in a predetermined position enables the listener to identify the sounds, (should separate images occur), but at the cost of providing additional information to the subject. For example, even if the two sound sources combine to form a single fused image the listener can still make meaningful lateralization judgements based on the average position of the fused sound image. One would expect, however, that the slope of the lateralization curve under these conditions to be somewhat less than that for a single sound source. On the other hand if the two images are totally separable the slope of the lateralization curve should remain unchanged whether one or two images are used. Part of this investigation has been concerned with detecting changes in slope for sources which are expected to yield inseparable, or fused sound images.

2.4 Philosophy underlying the Definitions

The purpose of this discussion is not to attempt a logical justification of the proposed definitions - one cannot prove a definition - but rather to describe the underlying philosophy which led to their formulation. In so doing it is hoped that the reader will at least agree with the reasonableness of the proposed measure of separability.

Having postulated the existence of a sound image one might ask whether it is possible to measure the properties or attributes of the postulated sound image in any way. The answer is a qualified yes. Although a sound image is a purely subjective phenomenon there is nevertheless a very strong and consistent correlation between subjective judgements of certain properties of a sound image and the physical characteristics of the associated stimulus. It is this widespread agreement in the subjective assessments of different people on different occasions which not only lends support to the concept of a sound image, but also enables various attributes of the image to be measured.

The attributes of a sound image, however, can only be measured by means of comparisons in the subjective domain. For example, one might postulate that the attribute X is the subjective correlate of the physical variate x, and that the two are monotonically related. Although X cannot be measured directly, it is possible by means of subjective comparisons in X to accumulate data either in support of, or contradictory to, the initial assumptions. If the data and assumptions are not inconsistent there is no reason for rejecting the original postulate. Hence although the various attributes of a sound image are individually defined and measured accordingly, this approach does allow the experimenter a useful means of classifying and identifying classes of sounds. It is important, however, not to lose sight of the essential arbitrariness of each classification.

Since a sound image is characterized by several subjective co-ordinates it seems reasonable to postulate that two <u>separate</u> sound images should be characterized by two distinct sets of subjective attributes. Hence it is argued that if a listener can distinguish between two separate stimuli (presented simultaneously) on the basis of a difference in at least one subjective attribute then the stimuli have in fact been separated. This interpretation of separability clearly depends on the subjective attributes delineated prior to the experiment.

Of the many subjective attributes characterizing a difference between two sound images, spatial separation is one of the most important. As Cherry (1961, 1962, 1963) has pointed out, the fundamental problem in aural perception is the separation and creation of sound images from the great mixture of sound stimuli reaching the two ears and as a consequence of this separation the concept of space arises. The proposed measure of separability is concerned primarily with the spatial separability of sound images; it being argued that if spatial separation can be shown then image separation as such must also have taken place, i.e. " ... one image cannot be subjectively placed in a distinct position from another, unless these images have first been separated (as gestalten)." (Cherry and Bowles 1960). Having postulated a measure of separability and its method of measurement a study was carried out on the practical feasibility of the technique. The implications of some of the results in terms of a crosscorrelation model were also considered.

2.5 Summary

Definitions for the total separability and degree of separability of two sound images have been postulated and the philosophy underlying the definitions explained.

A second property, _discriminability_, has also been distinguished and its relevance to the proposed method of measuring separability discussed.

CHAPTER 3 EXPERIMENTAL TECHNIQUES AND EQUIPMENT

3.1 Details of Experimental Technique

3.1.1 On Headphone Listening

The relative advantages and disadvantages of headphone rather than free-space listening have been discussed in the preceding review. It was decided to use headphones largely because of the greater control afforded to the experimenter. There appeared to be no strong reasons against the use of headphones and it seemed reasonable to assume that those factors influencing image separation with headphone listening would also be of importance in the free-space situation.

The headphones used were of the circumaural type. These afforded a good deal of ambient noise reduction (approx. 40 dB) and at the same time reduced the variability of the transducer position between fittings of the headset. The snugly fitting ear cushion also reduced low frequency leakage. Even so, it was found that removing and replacing headphones had a noticeable effect on the results. This was not surprising considering that the experimental technique involved interaural time delays of the order of microseconds; an increase in the airpath between transducer and eardrum of as little as 0.01 inches would correspond to an additional delay of this order. With supra-aural headphones the variability between fittings of the headset would have been considerably worse.

3.1.2 Interaural Time Delay as the Physical Correlate of Image Position

With headphone listening the apparent position of the sound image appears to lie within the head. Under these conditions there are two basic physical correlates of image position; interaural time delay, δT, and interaural intensity difference, δI. It was decided to use time differences for several reasons:

a) The transmission characteristics of the ear are only approximately linear. With δT as a variable, any non-linear distortion would be common to both ears.

b) The variation of image position with δI appears to be a secondary effect. Several theories of hearing suggest that intensity differences are transformed into time differences in the neural system, (e.g. see David, Guttman and van Bergeijk, 1958).

c) The variate δT is directly related to the time-displacement parameter, τ, in the running crosscorrelation between the signals at each ear.

d) Periodic stimuli show distinctive patterns in the crosscorrelation function which would otherwise be lost if δI were to be used as the main variable.

e) The range of δT that can be used before fusion breaks down is considerably greater than that occurring in real life.

There are, however, two major disadvantages in using δT:

i) The practical difficulties encountered in providing sufficiently large delays with the required accuracy are enormous.

ii) δT ceases to be a variable of any importance for continuous tones above about 1500 c.p.s. If the envelope is modulated, however, time delays are of importance.

Voelcker (1961) has pointed out that a combination of δI and δT corresponds more closely to the real life situation. The use of a low-pass network rather than a delay line would eliminate the above two objections without introducing any of the major disadvantages associated with using a pure intensity difference. On the other hand there is an infinite number of possible low-pass networks which could be used and selecting any one type would be extremely arbitrary. Furthermore, it would not be possible to compare results with those of other workers who have almost invariably used pure time delay. Nevertheless, the idea of using a special set of linear, low-pass networks as the physical correlate of image position is worth further consideration.

3.1.3 Instructions to Subjects

Subjects were told to sit, face the front and make themselves comfortable. Once comfortable they were expected to maintain the adopted position as best they could for the duration of the test. In particular, head movements were to be kept to a minimum. This was considered necessary since Bowles (1960) has shown that turning the head sideways has a marked effect on image position. Subjects wearing glasses were asked to remove them so as not to interfere with the fitting of the circumaural headphones.

The nature of the test was described briefly although the method of selecting δT for the various presentations was not disclosed. The most important instructions were those relating to the type of response expected. These usually went as follows:

"You will hear two sounds. One is a

(...... type sound (e.g.male voice)) *
(   ) the other is a
(sound near the centre of your head)

(...... type sound (e.g. female voice)   )
(   ).
(sound away from the centre of your head)

---

* The appropriate description was used depending on the type of test. In those experiments where only one sound source was used the subject was required to lateralize the associated sound image about his medial plane.

Please indicate the lateral position of the second sound relative to (the first sound / the centre of your head) by means of the handswitch provided. Only one of two possible decisions is acceptable - 'left' or 'right'. If you are uncertain you must guess. After you have made your decision the sounds will be switched off and another pair of sounds presented. The process is then to be repeated."

After a test the subjects were asked to comment on their subjective impressions and these were noted.

3.1.4 Other Details of Test Procedure

A single presentation usually lasted about 5 to 10 seconds and a complete test about 20 to 30 minutes. It was found in a special "fatigue" test that a subject showed a distinct decrement in performance after about an hour of continuous testing.

Auditory thresholds were measured prior to a series of tests and the signal level for all stimuli combined was usually fixed at 55 to 60 dB above threshold. The method of threshold measurement followed the Up-and-Down technique using a continuous stimulus of the type to be presented in the following tests. Steps of 5 dB followed by steps of 1 dB in the vicinity of the threshold were employed. Pure tone, air conduction audiograms were taken for all the subjects and no hearing abnormalities were observed.

The subjects were usually graduate students although in one experiment two schoolboys and a lady secretary were used. In that particular series of tests a distinction was drawn between "experienced" and "inexperienced" subjects. Testing times were selected in a quasi-random fashion depending on the availability of the subjects. The most popular period of testing was in the early evening between 6 and 9 p.m. Position of the headphones (which earpiece to which ear) was selected at random using a table of random numbers. Other experimental parameters such as cable pairs, matched delay lines, amplifiers, attenuators, etc. were similarly interchanged at random.

3.2 The Delay Unit

3.2.1 Preliminary Considerations

The delay unit formed the nucleus of the equipment and is therefore described in some detail. Experimentally, just noticeable differences in interaural delay as low as 10 to 20 microseconds have been observed (see preceding review Chapter 1), while on the other hand it has been reported that interaural delays of up to 10 -20 milliseconds can be tolerated without a breakdown in fusion (Cherry and Taylor 1954). A delay unit which would provide this range of delay with the required resolution posed formidable practical difficulties. Several possible methods of achieving a time delay were considered:

a) Using a real transmission line; although extremely accurate with an exceptionally flat frequency response the maximum feasible delays are only of the order of microseconds.

b) Building a lumped constant approximation to the continuous transmission line; the degree of approximation can be made very small depending on the number, size and precision of the components used. Design considerations indicate, however, that bandwidth must be traded for overall delay for any given circuit configuration. For a bandwidth of 5000 c.p.s. a total delay of up to a millisecond is feasible.

c) By means of recording techniques; the audio signal is recorded onto magnetic film or other medium and is monitored after a suitable time interval. This method is most appropriate for very large delays. Below a millisecond, however, the mechanical problems encountered (tape stretch, physical size of record heads) are almost insurmountable - some excellently designed machines (Holmes and Dukes 1954, Bowles 1960) have achieved a resolution no greater than 50 -100 microseconds. A magnetic drum would obviate some of the difficulties encountered with tape, but on the other hand minor eccentricities of the drum (less than 0.0001 inches) could cause a disturbing amplitude modulation of the recorded signal. Pulse or frequency modulation is normally used with magnetic drum equipment. This is expensive. All recording techniques are subject to inadequacies of the recording medium, the most noticeable being a decreased signal-to-noise ratio (40 - 50 dB) and wow or flutter.

d) By using acoustic or ultrasonic delay lines; although large delays can be obtained (e.g. see Brockelsby, Palfreeman and Gibson 1963), the poor frequency response of the transducers and reflections within the medium make this type of delay unit unsuitable for providing interaural delays.

e) By sampling techniques; perhaps the most versatile method, but unfortunately it requires a good deal of elaborate equipment. Signals would have to be sampled, pulse-code modulated, delayed using say a magnetic drum or shift register, demodulated and finally the original waveform reconstructed.

Although the last approach approximates most closely to the ideal it was not attempted because of the prohibitive cost in both time and money. It would, however, be a worthwhile scheme if sufficient funds ever become available (£2000 - £3000). The method adopted was that of using a lumped constant approximation to the continuous transmission line. It was not expected that delays any greater than a millisecond would be required in the proposed experiments; although had the demand for a longer delay arisen, a modified version of the tape unit used by Sayers (1957) could have been employed.

3.2.2 Design and Construction of the Delay Lines

The design procedure followed the approach recommended by J.M. Lester (in Zeluff and Markus 1949). An m-derived network was used with m = 1.4 for optimum phase linearity, and with buffer sections of m = 0.6 for matching to a resistive load. Mutual coupling between adjacent coils was employed in order to achieve an m greater than unity. According to the design specifications deviations from linearity in the amplitude or phase response were expected to be within 3% over 95% of the pass band. Design details are to be found in Levitt (1964).

A total of four delay lines were constructed; a matched pair of "short" lines, each providing delays of up to 58 microseconds in steps of 5.8 microseconds and a pair of "long" lines each providing delays of up to 580 microseconds in steps of 58 microseconds. Push-button switching was used for speed and accuracy; it was also convenient for coupling the system to the data recording unit.

In trading bandwidth for overall time delay the following compromise values were selected; 5 kc bandwidth for the long delay line and 50 kc bandwidth for the short line. Equalizing networks were added to the long line in order to compensate for effects resulting from the proximity of the cut-off frequency. Separate networks were designed for each delay setting based on a specially devised nomograph. Components of 1% tolerance were used throughout.

In designing the delay lines an ideally terminated network was assumed, which of course was only an approximation. The exact analysis of an incorrectly terminated ladder network consisting of a finite number of sections is a tedious task, although several algorithms have been suggested (e.g. see Levitt 1960, 1961, Ream 1961). Several small improvements in the response of the delay lines were effected as a result of considering the limitations of a real network. For example, an additional resistive shunt arm was introduced in order to balance the effect of a finite resistance in the series inductance. The ultimate evaluation of the design, however, lay in the final calibration of the equipment and it was found that deviations from the ideal were only slightly worse than the expected design values. It is interesting to note that a method of analysis suggested by the author (Levitt 1962) has been of use in another application (Ream 1962, section 9.4.).

3.2.3 Calibration of the Delay Unit

The amplitude calibration was carried out with the lines in circuit; direct comparisons between input and output voltage being made on a single meter. The results for the worst case (maximum delay) are given in Figs. 3.1 and 3.2 for the long and short lines respectively.

The phase or delay characteristics were measured by two separate techniques. One by using an electronic counter and measuring the time interval between axis-crossings of the test sinusoid; the other by subtracting output from input and computing relative phase from the residual. The latter method, being essentially a null method, was extremely accurate (within 0.1^{o}) for very small phase shifts and where there was a negligible difference in signal amplitude between input and output. It was particularly suited for measurements on the short delay line. The ultimate accuracy of the counter method was limited by the reliability of the triggering threshold of the device. The error was found to be practically constant and almost entirely independent of

FIG. 3.1 : CHARACTERISTICS OF 580 MICROSECOND LINE

(at Maximum Delay Setting)

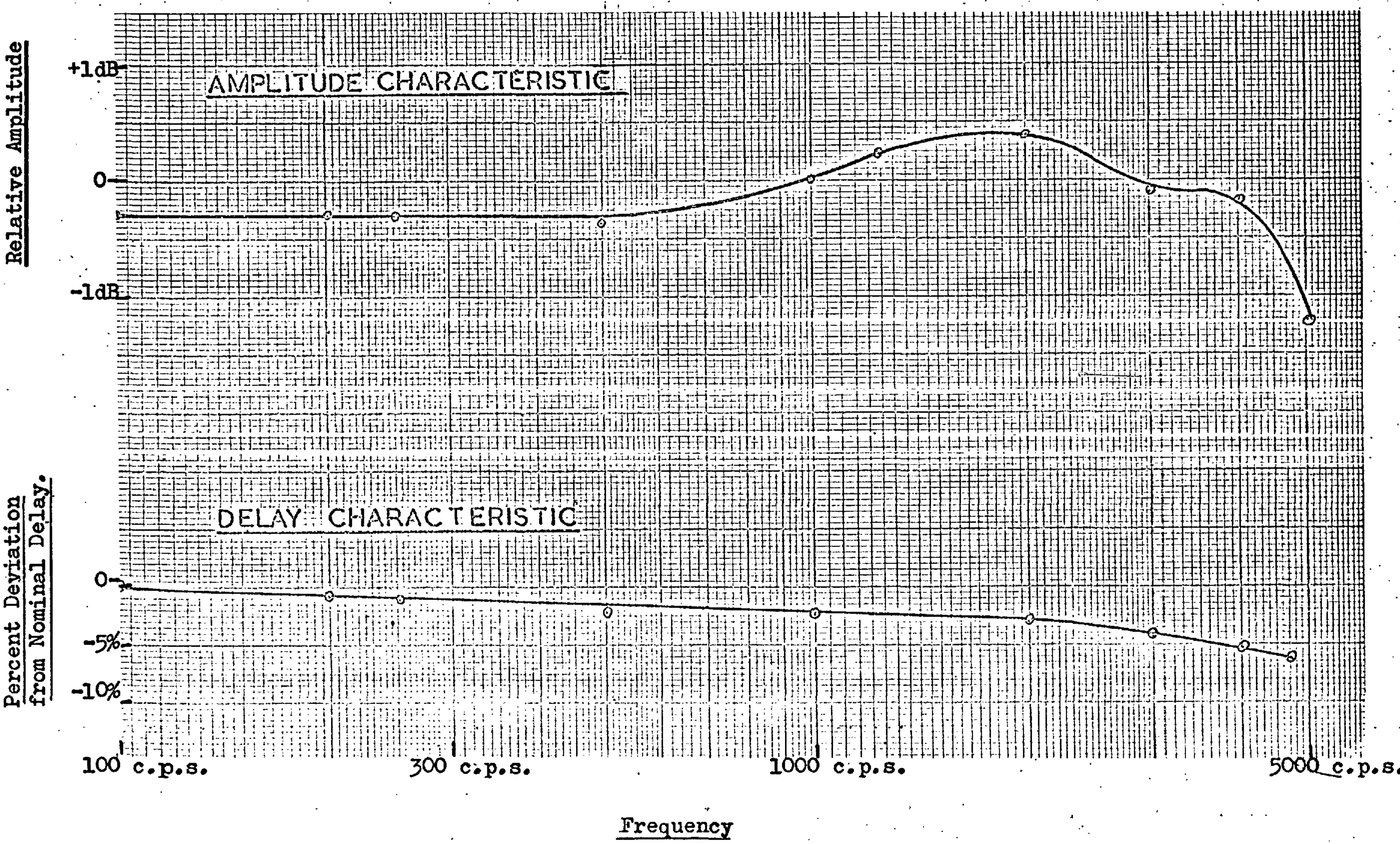
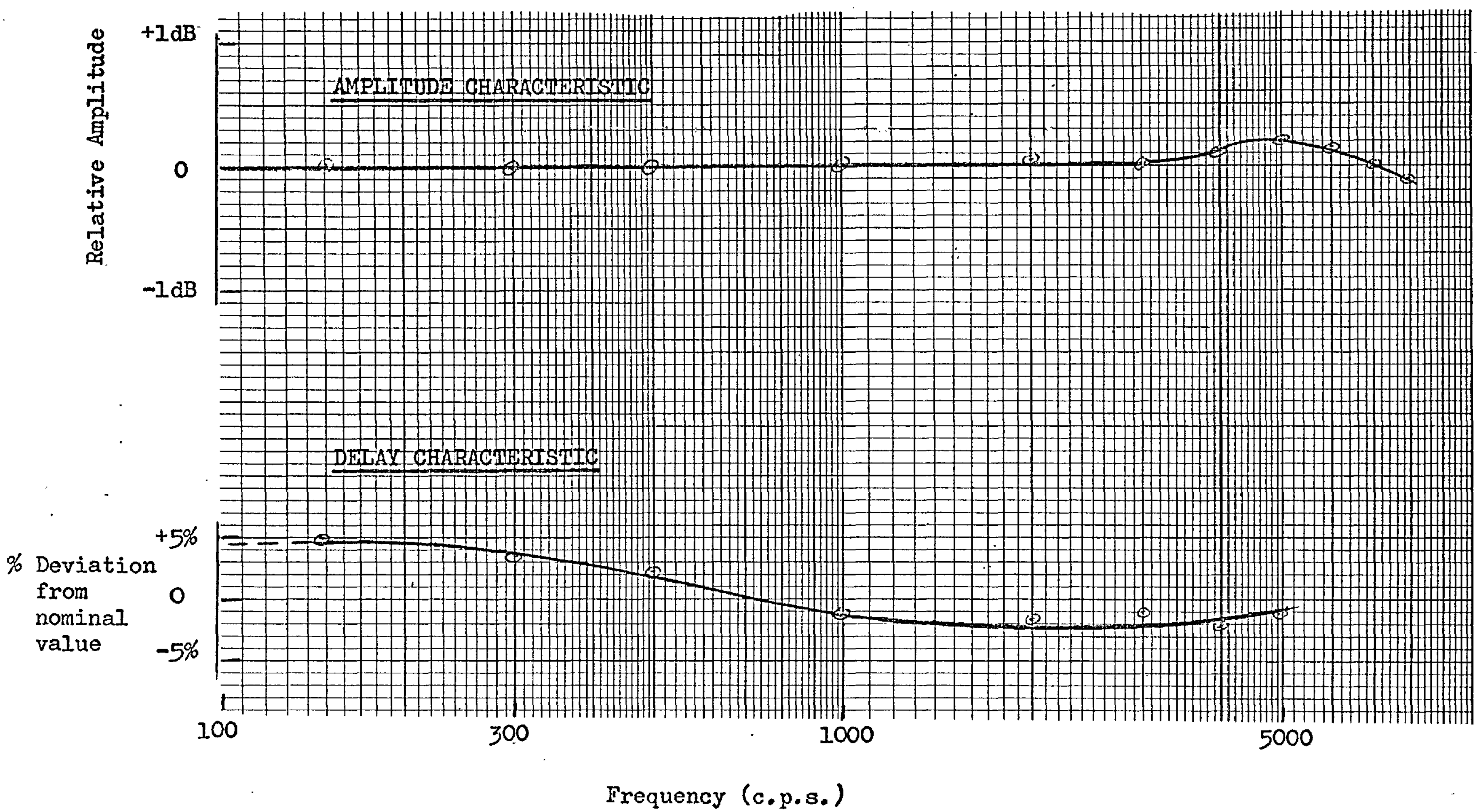


FIG. 3.2 : CHARACTERISTICS OF 58 MICROSECOND LINE

(at Maximum Delay Setting)



amplitude or delay characteristics. The relative error was therefore less for longer delays and was best suited for measurements on the long, 580 micro-second delay line (calibration accuracy: 0.1% to 1%). The delay characteristics for both lines are given in Figs. 3.1 and 3.2. Fig.3.7 shows the long line being calibrated using the counter method.

The delay measurements were checked using an Airmec type 206 phasemeter. This device measured the average of a series of rectangular pulses of constant height and of duration equal to the time interval between axis-crossings of the two signals. This was in some respects similar to the method using an electronic counter. Although slightly less accurate (resolution = 0.5 degrees), the phasemeter provided a useful check against gross errors. Good agreement was found between the various calibrations. Delay measurements using an electronic counter have recently received some publicity in the technical literature (e.g. McAleer 1961, Spencer 1963). The null method has been recommended by Wind (1956).

The measured values of delay for the different lines were found to be very close to their designed values. It is important to note, however, that the accuracy of any delay setting was only within 3 - 4%

of its nominal value; this was of particular importance for lines connected in tandem. It should also be noted that the frequency characteristics of the short line were considerably better than those of the long line. The cut-off frequency of the short line was above 50 kc and hence second-order effects were negligible.

The delay lines were designed as two matched pairs and a series of measurements were carried out on relative differences between the lines while in circuit. In these calibrations the extremely accurate null method was used throughout. There were no observable differences between the amplitude or delay characteristics of the two short lines. The precision of the amplitude measurements were within 0.2% (0 .02 dB); the delay measurements within 0.2 degrees. Greater precision would have been achieved were it not for non-linear distortions in the delay line (of the order of 0.2%). Measurements on the two long lines showed an equality between the two amplitude characteristics of within 0.1 dB and between the delay characteristics of up to 0.5%. The differences in delay for the two long lines are quoted in Table 3.1.

A significant characteristic of the delay lines was that whereas the nominal delay settings were only accurate within about 3 - 4%, the lines were extremely well matched to within a fraction of a percent in both amplitude and delay characteristics. In much of the experimental work it was the difference between delay lines which was of importance and not the absolute values. Only when a sound image was placed near the medial plane (i.e. interaural delay nearly zero) was the absolute value of delay important and in this case the short delay line with its markedly superior frequency and amplitude characteristics was used.

The effect on apparent image position of the ripple in the amplitude characteristic of the long delay line was difficult to assess. To a crude approximation a very small change in δI can be replaced by an equivalent change in δT. The trading relation between δT and δI, however, is extremely complex. A typical value appears to be about 0.2 dB/microsec. (David, Guttman and van Bergeijk 1958), which for the largest peak in the amplitude characteristic of the 580 microsecond delay line corresponds to a time difference of 2 - 3 microseconds. This could be interpreted as an additional non-linearity in the delay characteristic of about 0.5%.

TABLE 3.1 : DIFFERENCES IN DELAY BETWEEN 580 MICROSECOND LINES

Delay setting at maximum value.

Frequency(c.p.s.)	150	300	500	1000	2000	3000	4000	5000
Delay A-Delay B (micro secs.)	23	7	4	0	-2	-2	-4	-4
% Deviation	3.8	1.1	0.7	0	-0.4	-0.4	-0.7	-0.7

TABLE 3.2 : DIFFERENCES IN DELAY CHARACTERISTIC BETWEEN LEFT- AND RIGHT-HAND HEADPHONES

Frequency(c.p.s.)	150	300	500	1000	2000	3000	4000	5000
Delay Delay Right-Left (micro secs.)	31	0	6	13	11	4	-3	0
Estimated Standard Error (microsecs.)	22.2	7.5	4.5	2.2	1.2	1.0	1.9	1.3

Note: The above data are the means of 5 separate calibrations.

3.3 Headphones

Good quality moving coil headphones were used which provided a reasonably flat response over the frequency range of interest (up to 5 kc). Sharpe HA10 circum-aural headphones were selected for the preliminary experiments mainly because of their price and availability; in the major series of experiments circumaural headphones of special design were used. Ideally, condenser headphones would have been preferred but these were prohibitively expensive. The Sharpe equipment was reasonably good, but each headphone displayed a curious resonance in the 3 to 4 kc region. The effect on the apparent position of a pure tone in the vicinity of the resonance was quite marked.

The custom built headphones (Fig.3.4) were developed as a joint project with a colleague, R.E.C.White, and with advice from Edith L. Corliss of the National Bureau of Standards (Washington D.C., U.S.A.). The standard TDH39 cartridge was used. It was mounted in a moulded plastic shell* with hydraulic cushions containing glycerine, on the outside rim. Padding within the shell consisted of expanded plastic cut to the correct

---

* manufactured to specification by the Anticoustic Company, Guildford.

Fig. 3.3 : Flat Plate Coupler

Fig.3.4:Calibration
of the Special
Laboratory Headphones.

contour using an ingenious method invented by D.Harbison.*

The headphones were calibrated using a flat plate coupler as recommended by Corliss (1962) and Shaw (1962). The coupler consisted of a flat bronze plate ½" thick and 6" diameter, which could be mounted on the standard Brüel and Kjaer artificial ear. When in position the microphone diaphragm lay flush in the plane of the coupler, see Figs. 3.3 and 3.4. The headphone was placed over the microphone and held in position by a normal force of 800 grams in addition to its own weight. Vaseline was used to form an air-tight seal between microphone cartridge and coupler and also between headphone cushion and coupler.

Recordings of amplitude-frequency response were made using standard Brüel and Kjaer equipment. Calibrations were usually carried out at night during quiet periods, but even so, low-frequency vibrations in the building and other forms of pickup necessitated the use of a high-pass filter, (Fc = 100 c.p.s.). Differences in phase response between the headphones (Table 3.2) were measured using the Airmec type 206 phasemeter. Direct phase response measurements were confounded with

---

* of the Instrument Workshop, Department of Electrical Engineering, Imperial College of Science and Technology.

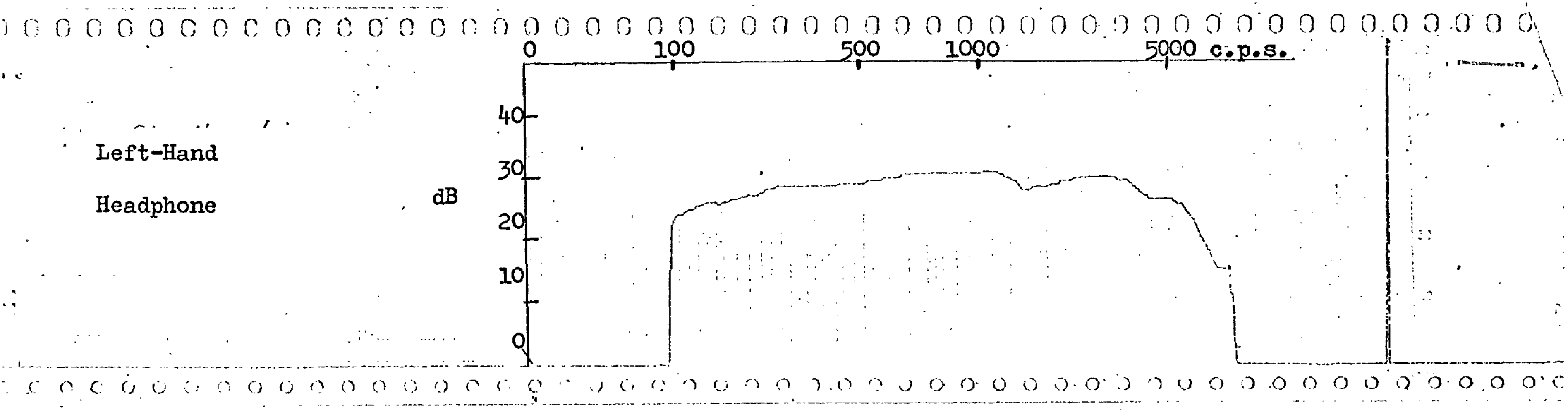
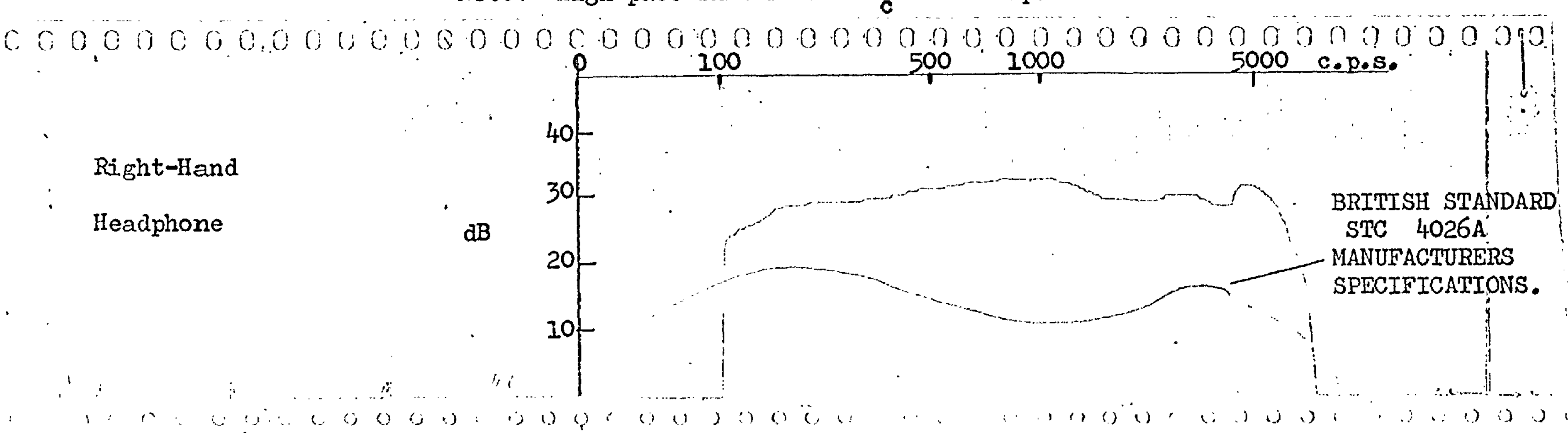


FIG. 3.5 : AMPLITUDE CHARACTERISTICS OF LABORATORY HEADPHONES

Note: High pass filter with F_c = 100 c.p.s. has been used.



the phase characteristics of the measuring equipment and hence these results have not been quoted; moreover the differences between headphones were of prime importance. The headphones were reasonably well matched in both amplitude and phase response. The frequency characteristics also compared very favourably with those of the British standard headphone, S.T.C. type 4026A, (see Fig.3.5).

It is important to bear in mind the limitations of the artificial ear calibration. One can never predict exactly the response as would be measured at the eardrum since there are far too many complex variables involved. The purpose of artificial-ear calibrations is to provide a measurement which is a close approximation to the true response, but at the same time is subject to precise specification and is repeatable. Shaw (1962) has shown that for circumaural headphones the calibrated response using a flat plate coupler compares favourably with the real ear response using a probe microphone at the entrance to the ear canal. Furthermore, the variability of successive calibrations is small, as shown in Table 3.2.

3.4 Data Recording Unit

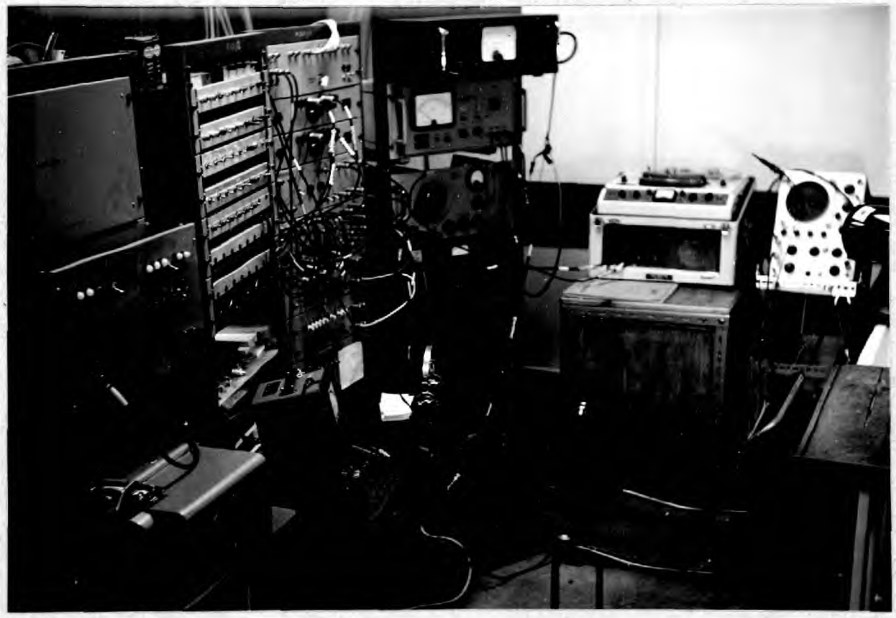
It was realized at the outset that the experimental programme would involve a good deal of data collection.

A single test usually required about 100 presentations and for each presentation it was necessary to record the values of δT for both sound sources and the corresponding response of the subject. Before the data unit was constructed this work was carried out by hand, which was both time consuming and liable to error.

The data unit (see Figs. 3.6 and 3.8) was designed to be as flexible as possible so as to be of general use. The device consisted essentially of a bank of 25 coding boards each of which had a total of 10 inputs. A decimal number could be stored by applying a 12 volt step to the corresponding input. Any desired computer symbol could be stored by applying a 12 volt step to a specified combination of these inputs; special plugs were wired for the most common symbols. On applying a control pulse the device punched onto standard computer tape the information contained at the inputs to the first n coding boards, where n could be set beforehand. After a second control pulse the next group of coding boards were scanned, and so on until all 25 coding boards had been scanned. The cycle could then be repeated. Provision was made for extending the number of coding boards to 50, or more, if desired.

In a typical experiment the various delay line settings and the subject's responses were each monitor-

Fig.3.6 : Layout of Equipment



Racks,(left to right): Data recording system, auxiliary rack, main control, generators and meters, tape recorder, oscilloscope.

Fig.3.7 : Calibration of Delay Lines

Equipment, left to right: Counter, high gain amplifier, delay line, oscilloscope.

Fig.3.8 : Rear of Data Recording Unit

Information is fed in through the bank of 25 plugs near the top of the unit.

ed to a separate coding board according to a prearranged sequence. At the end of a presentation all coding boards in the sequence were scanned. After an experiment a copy of the data tape was sent to the University Computer Unit for analysis. Several computer programmes were developed for sorting and analysing the data. Details of one of the programmes which has some novel features is given in Appendix A2. The probability of an undetected error in the data was found to be negligibly small ($< 10^{-8}$) owing to the high redundancy of the computer code; in addition, checks for possible misprints and inconsistent data were built into the computer programmes.

3.5 Ancillary Equipment

The remainder of the equipment was largely conventional consisting mainly of balanced pairs of amplifiers, mixers, switches, filters and attenuators. In order to reduce crosstalk separate d.c. supplies were used for left- and right-hand channels. The measured crosstalk figure was between 40 to 50 dB at various points in the system.

Since the delay lines were of limited bandwidth all signals were low-pass filtered on entering the system. Unless otherwise stated the low-pass filters were set at Fc = 4560 c.p.s. The rate of cut-off was about 130 dB/octave. Apart from the delay lines and filters the frequency response of the system was flat with negligible phase shift over a wide frequency range (20 to 20000 c.p.s.).

Stability of the apparatus was of great importance. Tests showed that the equipment could be reliably balanced to within 0.1 dB and this state of balance could be maintained almost indefinitely provided there was no marked change in ambient room temperature. Transistor circuits with considerable feedback were used throughout.

3.6 Concluding Remarks

The method of restricting the listener to a forced, binary choice has been shown to produce highly accurate and repeatable results. The testing procedure described in the early part of this chapter was a development of that used with great success by previous workers in the laboratory. A sensitive measuring technique, however, demands accurate equipment and this aspect of the apparatus was considered in detail.

The equipment centred on a delay unit which consisted of two pairs of matched delay lines. A lumped constant approximation to a continuous transmission line was used; it was necessary to trade bandwidth for overall delay and in the worst case the compromise values of 5000 c.p.s. bandwidth at 580 microseconds delay were used. An accurate calibration of the delay unit showed that although nominal delay settings were only accurate to within 3 - 4%, the two pairs of lines were extremely well matched, to within a fraction of a percent.

Commercially available headphones were found to be unsatisfactory and special circumaural headphones were developed in a joint project with a colleague. These compared very favourably in frequency response with the British standard headphone (S.T.C. Type 4026A).

Although every attempt had been made to balance the equipment, particularly between left- and right-hand channels, the matching was not perfect. The largest single source of unbalance was that due to the headphones. In carrying out the experiments headphones, channels, sound-sources and other matched sections of the apparatus were interchanged at random. Although randomisation is an extremely powerful tool for averaging out the effects of unwanted factors it would be unwise to rely too heavily on this artifice. In particular, it would be wrong to argue that since the effects of a lack of balance can be averaged out, great precision in the equipment is not justified. The cost of randomising out a significant factor is a much larger experiment; since psychoacoustic experiments are particularly time-consuming, time spent on equipment is usually well justified.

Considering that the experiments involved a good deal of data and speed of presentation was an important factor, a scheme for automatic data handling was devised and constructed. Apart from speeding up the experiments and automating the many tedious computations, the scheme was of value in eliminating human and other sources of error.

CHAPTER 4 STATISTICAL TECHNIQUES

4.1 Introduction

The purpose of this chapter is to discuss the characteristics and limitations of the statistical techniques relevant to this study. In particular, the problem of estimating difference limens, using the method of constants, is considered in detail.

The classic statistical tests (t-test, F-test, chi-square) were usually favoured, largely because of their proven reliability and efficiency. Similarly, standard experimental designs were used and these were generally of the factorial type. Although slightly less efficient than, for example, randomised block or Latin square designs, they were considered to be more reliable in that it was unnecessary to assume a lack of interaction between factors. The subsequent F-tests were also more robust. The use of a standard type of design enabled all the analysis-of-variance computations to be carried out on a digital computer (Mercury) using a single generalised programme (see Appendix A2).

Computer programmes were also developed for other routine statistical calculations, such as sorting the data and fitting the cumulative normal curve. The latter programme was based on an existing library routine developed by M. Treisman (1959).

4.2 On the Estimation of Difference Limens

4.2.1 Difference Limens and the Psychometric Function

Estimating the slope of the sigmoidal lateralization curve was the most important statistical problem arising in this investigation. In psychophysics generally, it is a problem of great importance. Given a psychometric function the associated difference limen, DL, may be defined as the probable error or standard deviation of the distribution (Guilford 1954, p.144, Torgersen 1958, p.141). The relationship between the 'average slope' of the psychometric function and the difference limen depends both on the exact form of the sigmoid and how 'average slope' is defined. In this study the sigmoidal lateralization curve has been assumed to be cumulative normal and the 'average slope', β, as $\sqrt{2\pi}$ times the slope at the midpoint. On the basis of these assumptions it turns out that $\beta = \frac{1}{\sigma}$, where σ is the standard deviation of the derived normal distribution and which in turn is defined as the difference limen, i.e. $\sigma = \text{DL}$. It should also be noted that with a cumulative normal the difference between the 30 and 70% points is very nearly σ, i.e. the difference limen.

4.2.2 The Cumulative Normal Assumption

In postulating a definite mathematical form for the psychometric function it is implicitly assumed that the function does in fact exist and that all the data are from a single regime, i.e. the measured values for a very large number of observations would all lie on a unique, smooth curve. The possibility that such might not be the case is considered in Section 5.5. In this discussion the assumption of a single regime for all the data is accepted implicitly. Sigmoidal distributions differ appreciably only at their extremities and since in this study interest centred on the crossover region of the curve it made very little difference which type of sigmoid was adopted. A cumulative normal was assumed since it was believed to be more appropriate theoretically (i.e. by invoking the Central Limit Theorem). On the other hand it might have been more practical from a statistical point of view to use the logistic function, or arc-sine transformation, (Eisenhart 1947, Berkson 1951, Maxwell 1959).

Having hypothesized a cumulative normal distribution it was found that the observed data were generally consistent with this assumption, except on those occasions where a drift in the listener's performance was observed. In these cases no fixed distribution could be assumed.

4.2.3 Probit Analysis

Converting to probits is equivalent to converting to a normal deviate plus 5. For the cumulative normal this transformation leads to a linear regression line; the addition of the constant 5 tends to eliminate negative probits for most practical investigations. Methods of fitting the probit regression line differ in the method of weighting (Irwin and Cheeseman 1939). Perhaps the most appropriate weighting is that which leads to the maximum likelihood estimates of the parameters defining the distribution. In this case the weights associated with each point on the regression line are inversely related to the expected variance in the probit domain. The practical difficulty with this approach is that it is necessary to know the position and slope of the regression line in order to determine the correct weighting. Irwin (1937) has suggested an iterative procedure in which the weights are estimated, the regression line fitted, the weights recalculated from the fitted regression line and the line fitted again; the process being repeated until two successive approximations agree sufficiently well. The estimates b and m thus obtained are asymptotically equivalent to the maximum likelihood estimates (Irwin and Cheeseman 1939, Garwood 1941). Tables and other aids for simplifying the computations have been developed by Finney (1947). The calculations are nevertheless rather laborious. Computational complexity was of little concern in this study since a digital computer was readily available; in the very early experiments, however, all calculations were carried out on a desk machine.

Maximum likelihood estimates have many desirable properties (e.g. see Mood 1950, p.158), one of the most noteworthy being that, asymptotically, they are most efficient. For small samples, however, this is not necessarily true. Furthermore, in the particular case of probit analysis difficulties arise with zero or 100% responses; the associated probit tends to infinity with weight approaching zero. Fisher has suggested using a fictitious probit based on the maximum likelihood estimate of the regression line. The method, however, is only approximate (Irwin and Cheeseman 1939), and the results need to be interpreted with care; moreover, if less than two batches provide responses other than zero or 100% then a probit estimate is not possible.

Although a cumulative normal transformation with differential weighting had been used by psychophysical investigators for many years, the probit transformation, as such, was only developed fairly recently (± 1930) for applications in the field of biology. The advantages of probit analysis lay in the emphasis placed on maximum likelihood estimation and in the comparative simplicity of the computations. Finney (1944) was the first to point out the value of probit analysis in subjective testing, while more recently Richards (1952) has shown an application of the technique in psychoacoustics.

4.2.4 Precision of Probit Estimates

Efficient estimation is generally of great practical importance. In this investigation it was particularly so since the duration of a test was limited and, for various reasons, it was not possible to combine the data of separate tests. Hence the ultimate precision of the technique was limited by the amount of information which could be extracted from a single test. Methods of reducing the sampling variance of the probit estimates were therefore considered in some detail.

Unfortunately, no unique solution to the problem could be found. Minimizing the sampling variance of m does not minimize the sampling variance of b. Furthermore the extent to which the sampling variance can be reduced depends on how much is known beforehand about the quantities to be estimated (e.g. see Friedman 1947).

Consider for example the estimated sampling variance of the maximum likelihood estimates (from Finney 1947);

$$V(m) = \frac{1}{b^2} \quad \frac{1}{Snw} + \frac{(m-\bar{x})^2}{Snw\ (x-\bar{x})^2} \qquad \ldots \quad 4.1$$

$$\doteq \frac{1}{b^2} \quad \frac{1}{Snw} \qquad \ldots \quad 4.2$$

$$V(b) = \frac{1}{Snw\ (x-\bar{x})^2} \qquad \ldots \quad 4.3$$

where m = midpoint of fitted regression line

b = slope of fitted regression line

x = main experimental variable (= δT for lateralization curve)

$\bar{x}$ = weighted mean = Snwx/Snw

w = weight attached to probit

n = no. of presentations for a given value of x, i.e. the number of observations per <u>batch</u>

S indicates summation, V(...) indicates "variance of ..."

Note that these formulae only hold for large samples.

The experimenter has complete freedom in the choice of x and n; furthermore w is related to the probit values as shown in Fig.4.1. Assuming that the cost of a presentation is constant, (i.e. independent of x) then the essential problem in planning a test

FIG. 4.1 : VARIATION IN WEIGHTING WITH EXPECTED PROBIT VALUE

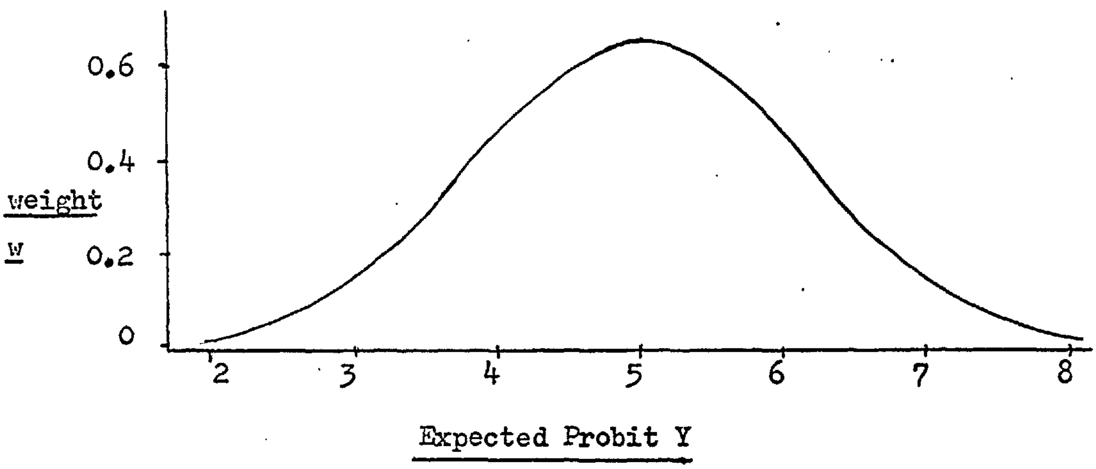
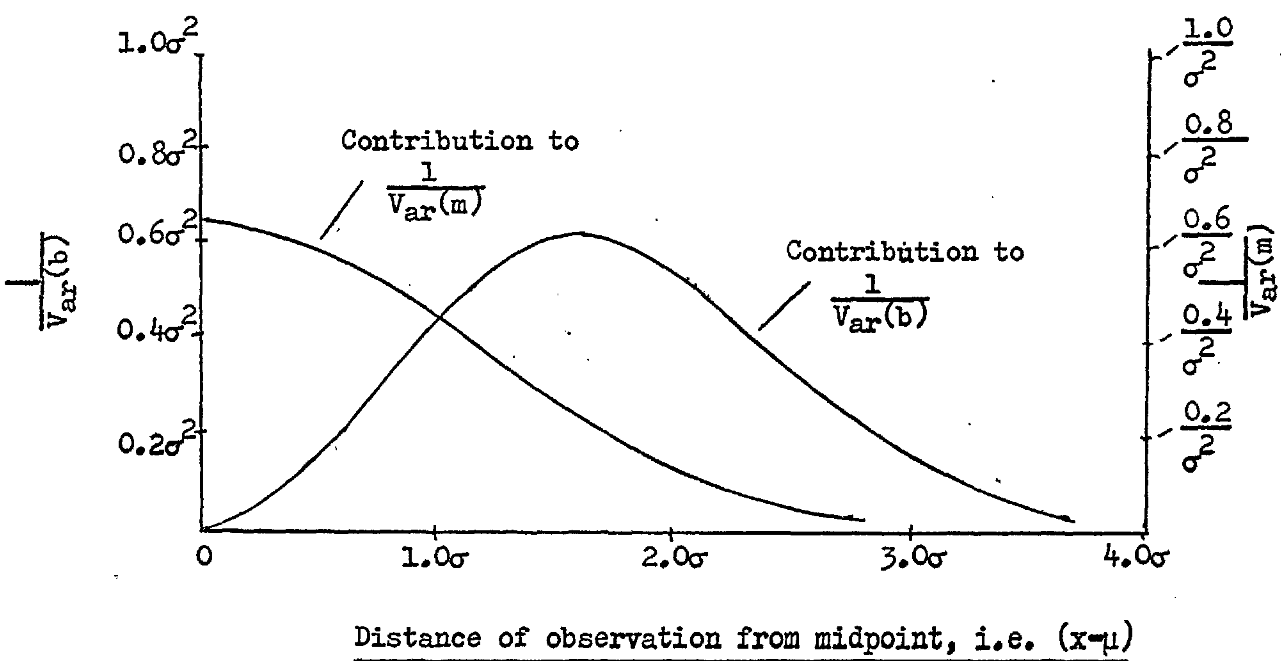


FIG. 4.2 : CONTRIBUTION TO RECIPROCAL OF VARIANCE

(for one of a pair of observations placed at a distance x on either side of the midpoint).



reduces to finding those values of x (i.e. δT) which will minimize either V(m), or V(b), or a combination of both. In addition it would also be of value if the resulting sampling variances were to be independent of the estimates themselves.

In a practical test it is recommended that the values of x_i be selected symmetrically about the true midpoint μ. This gives greater force to the subsequent goodness-of-fit test and also tends to reduce V(m). Assuming $\bar{x}$ and μ to be coincident, the contribution to $^1/V(b)$ and $^1/V(m)$ for a single presentation is shown in Fig.4.2. The maximum contribution to $^1/V(m)$ occurs at $x = \mu$ and to $^1/V(b)$ at $|x \pm \mu| = 1.57\sigma$. One cannot always ensure that $\bar{x}$ and μ are coincident; in some cases the purpose of the test may be in fact to estimate μ. Similarly a knowledge of β and μ is required in order to minimize V(b).

If little is known about β and μ then a good practical approach is to use batches of equal size and equispaced on the x-axis. This ensures that at least a few of the observations lie in an optimum position and also enables a useful goodness-of-fit test to be carried out subsequently. In addition, the overall contribution to $^1/V(b)$ and $^1/V(m)$ is reasonably independent of both β and μ since a smaller contribution to the reciprocal of the variance by the data of one batch is compensated for by an increased contribution by the data of another. The requirements of stability and efficiency conflict to a certain extent. The compromise is reflected in the number of batches used; fewer batches can result in greater efficiency but with a corresponding loss in stability. The use of five batches appears to be a useful compromise - it is a reasonably small number yet results in a comparatively stable variance and also allows for a meaningful goodness-of-fit test.

The possibility of stabilizing the variance by means of a transformation was also considered. This is clearly a complex problem owing to the large number of independent variables (n_i, x_i). For practical purposes, however, several useful restrictions may be introduced:

a) the x_i are symmetrical about the midpoint,

b) the x_i are equally spaced,

c) there are an equal number of presentations per batch, i.e. n_i = a constant, n.

In addition, the weighting function (Fig.4.1.) may be approximated by a parabola over a wide range. Using this artifice the expressions for the estimated sampling variances reduce to

$$V(b) \doteqdot \frac{K_1}{n\left[1-K_2 b^2\right]} \qquad \ldots 4.4$$

and $$V(m) \doteqdot \frac{K_3}{nb^2\left[1-K_4\ b^2\right]} \qquad \ldots\ 4.5$$

where the K_i are constants.

The approximation holds within about 5% over the probit range 3.5 to 6.5, which corresponds to the percentage range 7% to 93%.

It is important to note that whereas V(m) is independent of m it is critically dependent on b; hence estimates of the midpoint for probit regression lines of different slope must be handled with care. From (4.4) it would appear that there may exist a convenient transformation for stabilizing V(b); it turns out, however, that the assumption of a parabola for the weighting function breaks down at those values of b where the transformation is most needed. In addition, it is possible to choose an experimental design such that the constants K_1 and K_2 swamp the effect of b over the range of interest. The possibility of a stabilizing transformation was therefore not pursued further.

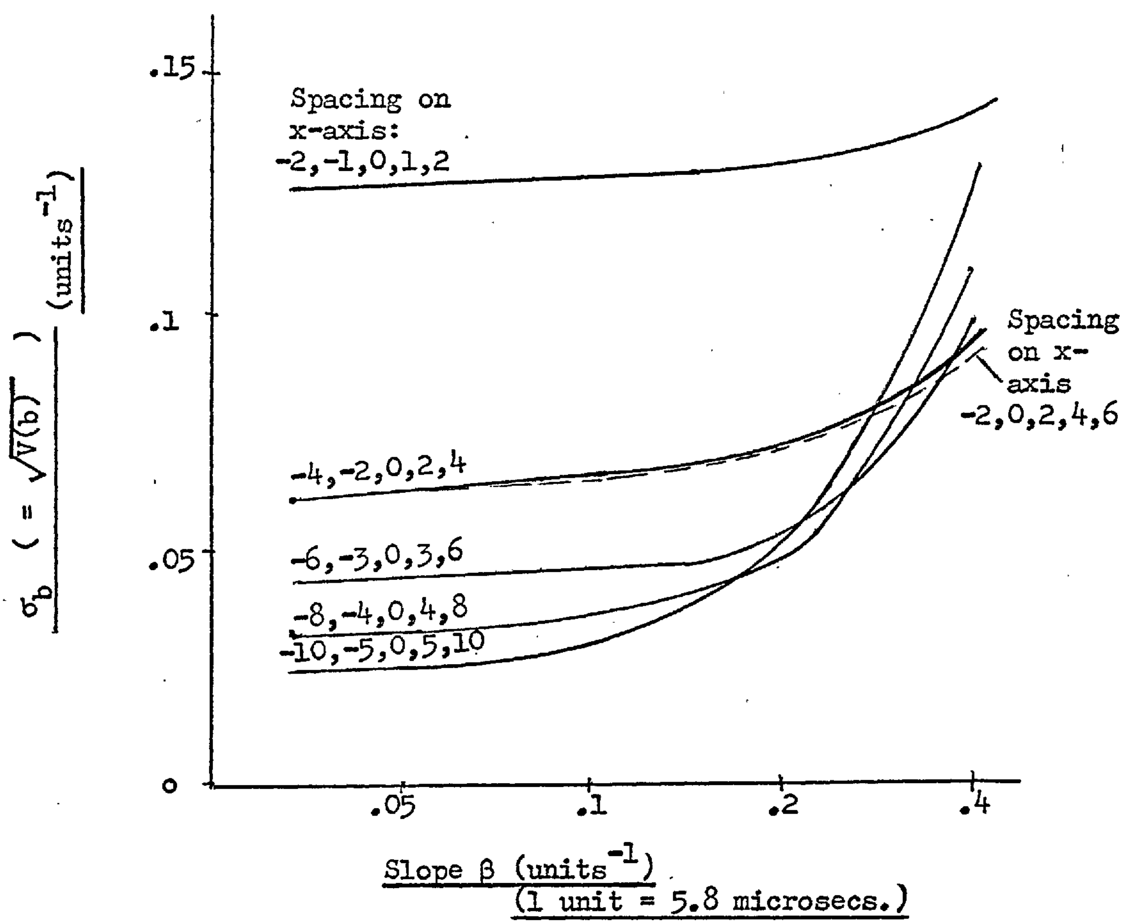
4.2.5 Towards an Efficient and Stable Test Design

In the experimental work interest centred largely on estimating and comparing the slopes of different probit regression lines. This required both efficient and stable estimates of β. Clearly it was necessary to establish an efficient practical design for this purpose. The advantages of using 5 equally spaced batches of equal size and symmetrically placed about the midpoint have already been discussed. The optimum number of observations per batch and the spacing of the x_i, however, need to be determined.

Since in the proposed design n is the same for each batch, both V(b) and V(m) are inversely proportional to n (eqns. 4.4 and 4.5). In a practical experiment, however, increasing n indefinitely will not necessarily reduce the error variance correspondingly since the error variance consists of both sampling and other sources of error. For example, in a complex design randomizing out the effects of different headphones may make a sizable contribution to the error variance. Furthermore, a large n may require too long a test resulting in fatigue and other factors detrimental to the listener's performance. The best value of n needs to be determined empirically for a given experimental situation. This was done in the first series of experiments (Series A) and a value of n = 10 was obtained.

Judicious spacing of the batches can lead to a stable estimate of β. In Fig.4.3. the estimated standard error ($=\sqrt{V(b)}$) has been plotted against β for a range of spacing intervals, (for a 5-batch, equispaced design symmetrical about the midpoint). A value of

FIG. 4.3 : STANDARD ERROR OF SLOPE ESTIMATE FOR VARIOUS SPACINGS OF δT ($\mu = 0$, $n = 10$)



$n = 10$ was used without loss of generality. Figures derived from Bowles (1960) and the results of preliminary experiments suggested a range of β from zero to 0.5 and hence a batch spacing of 2 units was considered to be most appropriate. As shown in Fig.4.3 this spacing yields a reasonably efficient and stable estimate of β over the range of interest. Furthermore, the assumption of $\overline{x} = \mu$ is not critical for this design as shown by the stability of V(b) when $\overline{x}$ and μ differ by as much as 2 units (see dotted curve, Fig.4.3). V(m) on the other hand is critically dependent on slope but is independent of μ as shown by (4.5). Care should therefore be taken whenever comparing estimates of the midpoint to check whether the probit regression lines are parallel, and if not, whether there is any correlation between m and b.

In view of the above the 5 batch, equispaced design with increments of 2 units on the x (i.e. δT) axis was considered to be the most useful. In the very first series of tests 9 batches at a spacing of 1 unit were used since at that time very little was known about the shape or size of the lateralization curve.

4.2.6 Estimation Bias

The maximum likelihood estimates have many desirable properties but they are not, in general, unbiased estimates. In particular, the maximum likelihood estimates of β for small samples show a marked bias. The extent of this bias can be seen in Fig.5.3 where the mean value of b over all factors in a typical experiment has been plotted against the number of presentations in a batch, n; i.e. q estimates of β each with n presentations/batch were rearranged to provide 2q estimates of β with $^{n}/2$ presentations/batch and 4q estimates of β with $^{n}/4$ presentations/batch and so on. The data of experiments A1 and A2 were used.

The bias appeared to be greater than 5% for samples of less than 10 presentations/batch. There was also a sharp rise in bias for very small samples of less than 5 presentations/batch; in these cases adjacent batches were combined as proposed by Finney (1947). The frequency of occurrence of zero or 100% responses was also very much higher in the small batches. A similar bias in slope estimation has been shown empirically by Berkson (1955) in sampling trials on a logistic function using three separate estimation techniques (maximum likelihood, minimum chi-square and minimum logit). Estimates of position, however, appeared to be unbiased provided the observations were symmetrically placed about μ. It should also be noted that the expressions quoted from Finney (1947) for the estimated sampling variance of m and b are only true for large samples.

4.3 Sequential Estimation Techniques

4.3.1 The Need for a Sequential Strategy

As pointed out in section 4.2.4, in order to estimate μ or β efficiently it is necessary to know their approximate values beforehand. Generally in planning any series of experiments some prior knowledge of the quantities to be measured is required; if this is not available, a pilot test of some sort is usually carried out. In essence this is a sequential strategy. What is here proposed is that a sequential strategy be used within a test. Clearly it would be of great advantage to use the knowledge of μ and β gleaned from previous observations in planning subsequent presentations.

Sequential methods involving quantal response data were developed during the last war largely for weapon testing purposes. Their application to psychological testing appears to have been first suggested by Dixon and Mood (1948); several other workers have since called attention to this approach (McCarthy 1949, Brownlee, Hodges and Rosenblatt 1953, Cornsweet 1962, among others). Much of the published work in this field has been concerned with the asymptotic properties of the estimates and not with the actual performance of finite samples. In addition, the emphasis has always been on estimating the midpoint of the curve, although estimates of other percentage points may be equally as important. A recent paper by Wetherill (1963a) is of importance since he considers both these problems, albeit for the case of the logistic function.

According to Wetherill it would appear that the slope of the logistic function cannot be satisfactorily estimated from sequential strategies. He found a large bias in the slope estimate for all the strategies he considered. In view of the great similarity between the probit and logistic curves it is reasonable to expect sequential strategies to be similarly inadequate for estimating the slope of the probit function. This is a disappointing result. On the other hand, single percentage points can be estimated very reliably, and given any two points on a linear regression line it is possible to derive the slope. For example, in the cumulative normal the difference between the x_i corresponding to the 30% and 70% points is approximately $1/\beta$ or σ, i.e. $x_{70} - x_{30} \doteqdot 1.05\sigma$. A very efficient and extremely practical strategy has been developed by Wetherill (1963b) for estimating any given percentage point over a wide range ($5\% < p < 95\%$). This strategy was found to have numerous advantages - both statistical and psychophysical - and it was adopted for the final series of tests; this being the first known practical application of the technique.

4.3.2 Wetherill's Sequential Strategy

In the simple Up-and-Down strategy (Dixon and Mood 1948) the experimenter adjusts the stimulus for each presentation in a manner which would tend to elicit a response opposite to that previously obtained. For example, an initial value of x reasonably close to μ would be used. If the response is negative x would be increased in steps until eventually a positive response is obtained; x would then be decreased in steps until a negative response is again obtained, and so on. The values of x used in the process oscillate about μ and hence μ can be estimated; the amplitude of the oscillations may be used for estimating β.

For estimating points other than the 50% point, Wetherill suggests the following strategy. Assume that the value of x associated with the percentage point p is to be estimated, i.e. x_p, where $p > 0.5$. A stimulus with x reasonably close to x_p is presented. If the response is negative then x is increased for the next presentation, but if the response is positive, x is not changed. Only after a series of N consecutive positive responses has been obtained is x decreased and the cycle once more begins again. The probability of a decrement in x is p^N; the probability of an increment is $1 - p^N$. The values of x used in a test will therefore oscillate about the point x_p where $p^N = 1-p^N = 0.5$. It is recommended that initially, regular steps of about $\frac{1}{2}\sigma$ be used. The spacing between levels may be reduced after at least six changes of direction in x (i.e. from increasing x to decreasing x and vice-versa). The set of x values between two consecutive changes is known as a run. The midpoint of a run provides an estimate of x_p; a more reliable estimate is the mean of several consecutive runs, i.e. the mean of an even number of changes. Such an estimate will be referred to as a Wetherill estimate. Fig.6.6 shows how the 30 and 70% points were obtained in an actual experiment. Note that $(0.7)^2 \doteq 0.5$, i.e. N = 2 was used.

It has been shown empirically (Wetherill 1963b) that if the spacing in x is less than σ the Wetherill estimates are more, if not as efficient as the maximum likelihood estimates - for samples of finite size. Adjacent runs were found to be highly correlated, but the estimates of every second run showed a correlation of less than 0.1.

An alternative estimation technique would be to use a sequential strategy for selecting the most appropriate values of x and then to use the method of maximum likelihood, or other procedures, for estimating the parameters of interest. Since the Wetherill estimates are highly efficient and easier to obtain there appears to be little advantage in using this approach.

4.3.3 The Use of a Sequential Strategy in Psychophysics

Apart from the greater efficiency in estimation, which is of considerable importance, a good sequential strategy has numerous other advantages.

a) The estimates are quickly and easily obtained and, if necessary, during the course of a test.

b) A faulty or meaningless experiment can often be detected at an early stage.

c) Long term variations in the parameter being estimated can be tracked during the course of a test, e.g. a subject's performance could alter during the test. It may also be possible to estimate when the change occurred.

d) The test may be discontinued once the desired precision has been achieved. This could be of great value in stabilizing sampling variances and also ensuring that the data is not unnecessarily precise;

e.g. other sources of error in a complex experiment may be considerably larger than the sampling error, in which case there would be little point in minimizing a small component of the overall error variance.

The following advantages, moreover, are of particular relevance to psychophysics.

e) The average difficulty of a test can be matched to the ability of the subject.

f) A randomized presentation sequence over a wide range of stimulus values is not always desirable (Miller and Garner 1944) and a sequential strategy can, if necessary, concentrate all presentations within a narrow region of the psychometric function.

g) In certain cases where a subject's performance is highly dependent on previous stimulation a standardized sequential procedure may prove to be more reliable than one where stimulus values are chosen at random. If an element of randomness is required this too can be arranged within the context of a sequential strategy (discussed later in this section).

h) ~~If the psychometric function is cumulative normal, or very nearly so, then those~~ Those percentage points which are most easily estimated using the Wetherill strategy are:

$$x_{50} \quad \text{(i.e. the midpoint)} \quad \text{for} \quad N = 1$$
$$x_{29.3} \quad \text{and} \quad x_{70.7} \quad \text{for} \quad N = 2$$
$$x_{20.6} \quad \text{and} \quad x_{79.4} \quad \text{for} \quad N = 3$$
$$x_{15.9} \quad \text{and} \quad x_{84.1} \quad \text{for} \quad N = 4.$$

That is, simple versions of the strategy may be used for estimating commonly used percentage points. (For practical purposes these are 20%, 30%, 50%, 70% and 80%). Furthermore, if the psychometric function is cumulative normal,

$$x_{70.7} - x_{29.3} \doteqdot 1.09\sigma \doteqdot \sigma$$

or, better still,

$$x_{84.1} - x_{15.9} \doteqdot 1.9964\sigma \doteqdot 2\sigma \quad .$$

Hence the difference limen, defined as σ, can be conveniently estimated using a simple version of Wetherill's strategy (i.e. $N = 2$ or $N = 4$).

It may also be noted that the weighting associated with x_{16} and x_{84} is an extremely good compromise for the probit estimates of both β and μ (see Fig.4.2). Hence this simple strategy is also of great value in placing the observations.

In fitting a regression line to only two points greater reliance is placed on the assumption of a specific curve for the psychometric function. If necessary, however, the sequential strategy could be used for estimating several points on the curve thus allowing for a meaningful goodness-of-fit test.

There is, however, a distinct danger associated with the use of sequential strategies in psychophysical experiments. Since the presentations follow a definite sequence it may be possible for the subject to detect a pattern and adjust his responses accordingly; either deliberately (malingering) or inadvertently. In the latter case a response characteristic of the strategy and not the psychophysical phenomenon may be observed. For example, a type of hysteresis effect could occur in which a different result would be obtained depending on the direction from which the p% point is approached.

These disadvantages can be avoided to a large extent by introducing an element of randomness into the strategy. For example, the x_i could be selected at random, but with fixed probabilities. Stochastic sampling schemes, however, converge comparatively slowly and are generally less efficient for small samples then the deterministic sequential strategy. A more practical method would be to employ two or more simultaneous strategies and to decide at random whether individual presentations belong to one or other strategy. Hence, although a deterministic strategy is used the successive values of x_i form a random sequence. The method of interleaving two strategies at random is particularly relevant to the measurement of difference limens, where two separate percentage points need to be estimated. This approach also has the advantage that, by selecting two complementary percentage points, the average frequency of positive responses in a test approaches fifty percent. Blackwell (1953) has pointed out that if relatively few positive responses occur in a test then the subject may exhibit a bias towards a negative response. It is also interesting to note that the idea of interleaving several simultaneous strategies has been recommended by Cornsweet (1962) for the simple Up-and-Down technique.

A practical comparison between the Wetherill strategy, the simple Up-and-Down method and the original quasi-random selection procedure was carried out, (Chapter 6). The results suggested that there might be a slight difference in a listener's performance depending on the strategy employed. This point, however, was difficult to establish since a drift in m might have occurred during the tests; an effect which is not catered for by the quasi-random technique.

4.4 Summary and Conclusions

The validity of the cumulative normal assumption was discussed; it being pointed out that any reasonable sigmoid would have sufficed since these curves differ appreciably only at their extremities. The method of placing observations for an efficient and stable estimate of slope was considered in detail and a practical experimental design was derived. This design was used in the early experiments. Later experiments, however, required the use of a sequential strategy and sequential methods were considered in some detail. A newly developed strategy, which may find useful application in psychophysics generally, was described and its advantages, both statistical and psychophysical, enumerated.

The major danger in using a sequential strategy is the possibility of inducing sequential effects in the subject's responses. This difficulty and the possibility of malingering by the subject can be largely avoided by using more than one strategy and interleaving presentations for each strategy at random.

CHAPTER 5 SERIES A; PRELIMINARY EXPERIMENTS

5.1 Introduction

This series of experiments was primarily designed to investigate the repeatability of the technique and to determine the statistical properties of the data. In particular it was of interest to find out how the proposed measure of separability was related to the inherent variability of the data, the variations between subjects, between tests and within a test. It was also of interest to determine the consistency of the many assumptions made in the statistical analysis.

Two statistically identical, random, white noise sources subsequently bandlimited to 4560 c.p.s. were used as the basic stimuli. The reason for wanting to measure the "degree of separability" of two noise sources was that this clearly represented an extreme case in lack of separateness. Two truly random, statistically identical noise sources which are superimposed should be theoretically indistinguishable from a similar, single noise source of twice the power - if this were not the case the sources could not be statistically identical. In order to ensure that the sources were in fact statistically identical, tape recordings were used where each recording was made under the same conditions and using the same noise generator. The recordings were usually made within minutes of each other.

The experimental technique was essentially that described in section 2.3 for two sounds which may possibly be separable but not discriminable. Since it was believed that differences between subsequent test periods (which involved removing and replacing headphones) might have an important effect on the results, a possible factor, occasions, was defined.

An occasion refers to that uninterrupted period of time during which the listener is tested. During an occasion the headphones were not moved and the listener kept his head and body position reasonably steady. Usually a single occasion consisted of a single test although sometimes more than one test was carried out during an occasion. The possibility of combining the data of several occasions to make up a single test was considered in the light of the experimental evidence (section 5.2.3).

The data within an occasion could be grouped according to their time sequence of occurrence, e.g. of the 200 presentations making up a test, the first 50 could form one group, the second 50 another group, and so on. Differences between groups were analysed as a time-order factor. Clearly it was desirable to have as many groups as possible since the larger the number of groups the greater the degrees of freedom and the more powerful the subsequent statistical tests; on the other hand, the smaller a group the larger the sampling error. The best compromise depended on the variability of the data and hence could ~~not~~ only be determined empirically. Part of this investigation was concerned with finding the most useful grouping of the data.

5.2 Experiments A1 and A2

5.2.1 Description

Experiments A1 and A2 were carried out concurrently. The purpose was to investigate the effect of subjects, occasions and time-order on the slope of the lateralization curve for one and for two noise sources. In this joint series nine batches of 20 presentations were used on each occasion. The batches were placed at equi-spaced intervals on the δT axis. A spacing of 1 unit (=5.8 microseconds) was used. The presentations followed a random time sequence as obtained by carefully shuffling a pack of cards in which each card contained the information for one presentation. The first 20 presentations were separate to the above and were used for the dual purpose of allowing the listener to settle down and enabling the experimenter to check that the range of δT chosen was more or less symmetrical about the position of the medial plane.

In experiment A1 a single, random, white noise source was used and the listener was required to lateralize the resulting sound image about his medial plane. The effect of the following factors on both b and m were considered:

Subjects (3 levels; I.E., G.S. and P.D.)

Occasions (4 levels)

Time-Order (4 levels).

The design was a simple factorial with 1 observation/cell. A fixed-effects model was assumed. The data and associated analysis of variance are given in Tables 5.1 to 5.4.

In experiment A2 either one or two noise sources were used; each was random, white noise subsequently bandlimited to 4560 c.p.s. When two sources were used one was placed as nearly as possible in the medial plane using a variation of the Up-and-Down method; this was done at the start of the test. The value of δT for the second source was varied at random. The subject usually claimed to hear one sound whether one or two sources were used; in all cases the subject lateralized the resulting sound about his medial plane. The effect on b and m of the following factors was investigated:

Subjects (5 levels; G.S., P.D., P.W., J.B. and K.P.)

TABLE 5.1 : EXPERIMENT A1, b ESTIMATES

Time-order 1	2	3	4	Subjects	Occasions
.520	.411	.314	.229	G.S.	1
.259	.167	.252	.325	I.E.	
.443	.592	.344	.215	P.D.	
.415	.403	.373	.501	G.S.	2
.150	.366	.294	.303	I.E.	
.281	.390	.195	.270	P.D.	
.250	.457	.347	.351	G.S.	3
.068	.186	.232	.036	I.E.	
.310	.507	.459	.340	P.D.	
.393	.196	.693	.286	G.S.	4
.478	.334	.395	.240	I.E.	
.199	.350	.370	.452	P.D.	

All values quoted in units of 5.8^{-1} microseconds.$^{-1}$

TABLE 5.2 : ANALYSIS OF VARIANCE IN b FOR A1

Effect		Mean Square	Significance Level
Time-Order	(A)	.0127	-
Subjects	(B)	.0735	approx. .99
Occasions	(C)	.0102	-
Interactions:	A x B	.0081	-
	B x C	.0215	-
	A x C	.0175	-
Residual S. of Sq.		.0124	-

Mean Values: Subjects	G.S.	I.E.	P.D.
(units as above)	.384	.255	.357

TABLE 5.3 : EXPERIMENT A1, m ESTIMATES

Time-order 1	2	3	4	Subjects	Occasions
-2.174	-2.979	-3.064	-3.837	G.S.	
-0.511	-1.985	-1.195	-0.349	I.E.	1
-3.964	-2.757	-4.250	-3.269	P.D.	
-1.131	-1.047	-0.388	-0.819	G.S.	
3.999	4.508	5.296	4.822	I.E.	2
-0.007	-0.455	-1.185	0.714	P.D.	
0.018	-0.187	0.190	-1.174	G.S.	
3.005	-0.295	2.352	8.631	I.E.	3
-1.958	0.359	-0.949	0.357	P.D.	
-0.892	-1.806	-2.716	-3.393	G.S.	
5.446	6.812	5.518	6.272	I.E.	4
1.986	-1.440	0.284	0.543	P.D.	

1 unit = 5.8 microseconds.

TABLE 5.4 : ANALYSIS OF VARIANCE IN m FOR A1

Effect		Mean Square	Significance Level
Time-Order	(A)	1.63	-
Subjects	(B)	112.47	.995
Occasions	(C)	41.06	.995
Interactions:	A x B	2.59	-
	B x C	7.19	.95
	A x C	1.30	-
Residual		1.86	

Mean Values: Subjects	G.S.	I.E.	P.D.	Occasions	1	2	3	4
(units as above)	-1.587	3.270	-0.999		-2.528	1.192	0.862	1.3[illegible]

TABLE 5.5 : EXPERIMENT A2, b ESTIMATES

Time-order 1	2	3	4	5	6	7	8	No. of Sources	Subjects
.428	.612	.414	.471	1.103	.408	.432	.382	1	G.S.
.043	.449	.416	.041	.116	.176	.167	.205	2	
.561	.639	1.006	.967	.268	.616	.626	.727	1	P.W.
.518	.391	.358	.196	.052	.653	-.043	.131	2	
1.069	.512	.344	.391	.456	.537	.574	.594	1	I.B.
.307	.426	.612	.267	.213	.125	.288	.354	2	
.359	.085	.208	.262	.457	.298	.144	.519	1	K.P.
0.39	.164	.171	-.038	.280	.267	.082	.250	2	
.313	.380	.509	.244	.192	1.174	.337	.692	1	P.D.
.203	.113	.557	.159	.179	.124	.023	.214	2	

All values in units of 5.8^{-1} microseconds.$^{-1}$

TABLE 5.6 : ANALYSIS OF VARIANCE IN b FOR A2

Effect		Mean Square	Significance Level
Time-Order	(A)	.049	-
No. of Sources	(D)	1.434	.995
Subjects	(B)	.136	.95
Interactions:	A x D	.028	-
	D x B	.051	-
	A x B	.044	-
Residual		.045	

Mean Values: No. of Sources	1	2	Subjects	G.S.	P.W.	I.B.	K.P.	P.D.
(units as above)	.508	.240		.366	.479	.442	.244	.338

TABLE 5.7 : EXPERIMENT A2, m ESTIMATES

Time-order								No.of	Subjects
1	2	3	4	5	6	7	8	Sources	
-.016	-2.80	0.34	-0.68	-0.83	-1.14	-0.21	-1.18	1	G.S.
-8.68	-0.98	-1.08	-5.66	-0.98	-2.29	-1.77	-2.51	2	
-1.34	-0.76	-1.12	-1.19	0.25	0.014	0.50	1.55	1	P.W.
0.33	-0.20	0.60	1.81	9.23	1.25	1.50	6.09	2	
-0.51	1.22	-0.78	0.48	0.09	-1.02	0.92	2.29	1	I.B.
0.13	-0.34	-2.80	-2.44	-0.55	-3.16	0.14	1.04	2	
-3.02	1.56	-3.03	-0.56	-0.31	-0.04	1.45	1.67	1	K.P.
3.97	-0.60	-0.24	-10.1	3.91	-0.36	3.50	-0.84	2	
2.03	0.97	-1.96	0.94	-0.97	-0.51	0.39	0.72	1	P.D.
1.32	0.35	-0.36	-2.72	3.58	-1.36	-16.5	1.09	2	

1 unit = 5.8 microseconds.

TABLE 5.8 : ANALYSIS OF VARIANCE IN m FOR A2

Effect		Mean Square	Significance Level
Time-Order	(A)	9.43	-
No. of Sources	(D)	9.30	-
Subjects	(B)	21.6	.95
Interactions:	A x D	8.78	-
	D x B	20.8	.9
	A x B	8.82	-
Residual		7.76	

Mean Values: Subjects	G.S.	P.W.	I.B.	K.P.	P.D.
(units as above)	-1.904	1.157	-0.332	-0.193	-1.256

1 or 2 Sources (2 levels)

Time-Order (8 levels).

As before, a factorial design assuming a fixed-effects model was used. Some of the data were common to both experiments. The data and associated analysis of variance are given in Tables 5.5 to 5.8.

5.2.2 Results

In both experiments and for both b and m it was found that differences between subjects were highly significant. On the other hand it would appear that time-ordering had an insignificant effect on both b amd m, although the results in Table 5.3 suggest that on certain occasions a slight drift in m might have occurred. The occasions factor had an interesting effect in that it appeared to be significant in m but not in b.

The most significant factor by far was that of using 1 or 2 sources. There was a marked decrement in estimated slope when two sources were used. The magnitude of the estimated difference in slope, however, was only about twice the standard error. Only two types of interaction were observed; between subjects and occasions in m and between number-of-sources and subjects, also in m. No significant interactions were observed in b.

It should be remembered that the sampling variance of m depended on b; the sampling variance of b, however, was expected to be relatively stable and uncorrelated with m (see section 4.2.4). The magnitude of the observed variations in m, however, were so large as to swamp any effects due to non-homogeneity of the sampling variance.

5.2.3 Discussion

The experimental results were encouraging in some respects, but disappointing in others. Perhaps the most outstanding feature of the results was their accuracy and repeatability. The estimated difference limen, $^1/b$, for a single noise image was found to be about 3 units on average (15 to 20 microseconds) with a standard error of approximately a third this value. (The DL was defined as the standard deviation, σ, of the derived normal distribution.) The remarkable sensitivity of the auditory system to interaural differences has been demonstrated by several investigators and the results here obtained compared favourably with previous work, (see review in Chapter 1). The small standard error was a good indication of how well the experimental design balanced out the effects of large extraneous factors; such as for example, differences between headphones. The effect of factors which varied at random within a test, i.e. between presentations, could not be randomised out, however, since these were implicitly confounded with the main effects. Stability of the equipment was therefore of great importance.

A disappointing aspect of the results was the small change observed in the slope of the lateralization curve for one and for two noise sources. The standard error of b was only about $\frac{1}{2}$ the estimated change in slope. Considering that two statistically identical, random, noise sources represent an extreme case, these results indicated that the range of the proposed measure of separability would be small. Of greater concern, however, was the supposed reliability of the technique used in placing an image in the medial plane. The results have shown the listener to be extremely sensitive to changes in δT, hence it would seem unlikely that in the preliminary run the fixed image could be placed in position with the required accuracy. As the tests with one and with two sources were performed on separate occasions and the estimated midpoint, m, varied widely between occasions it was not possible to check whether the position of the fixed image was sufficiently close to the position of the medial plane (as measured in terms of the physical variable δT). A subsequent experiment, (A3), was devised to check this point.

Another matter of great concern was the relative stability of the position of the medial plane. Even if it were possible to correctly place the "fixed image" at the start of a test, was it safe to assume that it would remain in the correct position throughout the test? Other workers (see Chapter 1) have demonstrated a marked shift in medial plane position by prolonged or unusual stimulation. In view of this and considering the general variability of human performance it did not seem unlikely that small drifts in the midpoint of the lateralization curve could occur during a test. The problem was to establish experimentally whether or not such drifts occurred. In this respect the apparent insignificance of the time-order factor was of great interest. It showed that at least if one arbitrarily classified the data into consecutive groups there appeared to be no significant change in slope or midpoint during an occasion. This, however, should not be regarded as sufficient evidence that no change at all occurred with time. Testing for a change in regime is a difficult problem if one has no knowledge of how or where the change could have occurred. For example, by arbitrarily classifying the data into groups the starting point of a drift may have been averaged out within one of the groups hence making the subsequent statistical test less powerful. In addition, the sampling error would have been increased by subdividing the data. The fact that variations in b between occasions were found to be insignificant suggested that for slope at least, changes within an occasion were unlikely. It was conceivable, however, that a gradual drift in the midpoint could have occurred with certain subjects on certain occasions. The possibility of such a drift has been investigated further in section 5.5.

Variations between subjects were found to be large in both m and b. This was not surprising since large differences between people have often been observed in psychophysical investigations. In view of these considerable variations and the difficulty of obtaining a truly representative sample of the human population, it was considered preferable to do a large and balanced set of tests on a few people rather than a few tests on many subjects. In this way a smaller error variance would be obtained but at the expense of a loss in generality. Considering the exploratory nature of these investigations this was not felt to be serious; moreover it seemed reasonable to assume that the results obtained with a small crew of listeners would be qualitatively, if not quantitatively, representative of the general population.

It is interesting to note that variations between occasions were significant in m but not in b. The variation in m was unfortunate in that it prevented the combining of data from separate occasions. The ultimate precision of the estimates therefore depended on how much information could be extracted during a single occasion. This accentuated the need for a highly efficient estimation technique. Considering that an air path of 0.01 inches involves a time interval of about 1 microsecond it would not be surprising if the observed variations in m between occasions were due largely to variations in the fitting of the headphones. Had supra-aural headphones been used the variability would most probably have been greater. Fixed differences between the headphones may have had some effect on m. This factor, being of nuisance value only, had been randomised out and hence its effect could not be evaluated directly. Rearranging the data, however, indicated that any such effect would be small compared to the random fluctuations in m between occasions, (see section 5.4).

Significant interactions between factors were few. In particular, none of the factors influencing b showed any interactions. This was a useful result in that those factors which affected lateralization ability appeared to act independently. It was also a result of some practical value in that it warranted the use of Latin square and other labour-saving designs. A marked interaction in m between subjects and occasions was observed. This was not surprising considering the magnitude of the two main effects; moreover it supported the belief in a complex interaction between subjects, occasions and stability of medial plane position. The observed interaction in m between subjects and the number of sources may possibly have been spurious in view of the fact that only one of the associated main effects (number of sources) was significant; furthermore, there were no reasons for expecting such an interaction to exist. Considering that over a dozen F-tests indicating no significant effect had been carried out, it was not unlikely that at least one would show a significance level of just under 0.95 purely by chance.

5.3 Experiment A3

5.3.1. Purpose

The preceding investigation raised the following important question; could the experimenter reliably place the fixed sound image in the medial plane. It was decided to check this by carrying out tests with one and two sources concurrently (by interleaving the presentations); if a significant difference in m between the two cases could be detected then this would be evidence that the fixed image had been incorrectly placed.

5.3.2. Description and Results

The technique was essentially the same as that used in experiments A1 and A2. Presentations using one or two sources were interleaved at random. The subject was instructed to lateralize the resulting sound about his medial plane. One subject was used on five separate occasions. The experiment was small in size since it was aimed at demonstrating a negative effect; furthermore, a marked change in m was expected. The factors investigated were

Number of sources	(2 levels)
Occasions	(5 levels)

with 2 replications per cell.

The results are quoted in Tables 5.9 and 5.10.

5.3.3 Discussion

Although a small experiment, the results showed a significant difference in m for one and for two sources. This consolidated the belief that, with the given test procedure, the fixed image could not be reliably placed in the medial plane. Some rethinking was obviously necessary and it was for this reason that the more efficient sequential strategies were first considered.

In addition, the listener reported that although a single image was heard, it exhibited a change in subjective quality in some presentations, as if the image had become wider. For experiments of this type it was clearly of importance to determine the sensitivity of a

TABLE 5.9 : EXPERIMENT A3, ESTIMATES m AND b.
- Subject G.S., 2 replications/cell.

b Estimates		m Estimates		No. of Sources	Occasions
.493	.852	-0.23	0.18	1	1
.191	.120	0.02	-0.56	2	
.644	.604	0.11	-1.10	1	2
.083	.050	-1.95	-6.37	2	
.676	.719	-1.23	-0.82	1	3
.065	.170	-7.50	-1.68	2	
.680	.567	-1.05	-0.79	1	4
.117	.118	-2.97	-4.61	2	
.461	.670	1.25	1.67	1	5
.130	.230	-2.31	0.51	2	
1 unit = 5.8^{-1} microsecs^{-1}.		1 unit = 5.8 microsecs.			

TABLE 5.10 : ANALYSIS OF VARIANCE, EXPT. A3

Effect	b Estimates		m Estimates	
	Mean Square	Significance Level	Mean Square	Significance Level
No. of Sources (D)	1.296	.995	32.2	.95
Occasions (C)	0.0033	-	8.16	nearly 0.9
Interaction C x D	0.0057	-	1.92	-
Error Sum of Squares	0.0109		3.32	

listener to changes in distance (along the δT-axis) between two statistically identical noise sources, and whether in fact two separate sound images could be heard if the spacing between sources was sufficiently large. This point was considered in detail in subsequent experiments.

Although the experiment was primarily concerned with determining m, the observed results for b were also considered. Variations in b were found to fit the pattern established by experiments A1 and A2; i.e. there was an extremely significant difference between using one and two sources, but that variations between occasions were negligible.

5.4 Subjective Tests for Lack of Balance in the Equipment

Although physical measurements had shown the equipment to be reasonably accurate, the ultimate check lay in a subjective test. Checks for possible malfunctioning of the equipment were therefore carried out whenever possible.

A matter of some concern was the possibility of a lack of balance between left- and right-hand channels which could be related to the position of the delay lines. This was checked during the previous series of experiments by switching the delay line into either channel at random for those presentations involving zero delay. A record was kept of delay line position and the associated response of the listener, (Table 5.11). The test showed no significant effect on the listener's responses either within individual tests or over all tests combined.

The possible effect on image position of differences between the headphones was also of interest. Although this factor, being of nuisance value only, had been randomised out it was possible by rearranging the data to get some idea of its effect. Table 5.12 shows the combined results of several experiments using the special laboratory headphones; since channels were also interchanged at random, this factor has been tabulated as well. The results showed an effect consistent with a fixed bias in the listener and no noticeable effect due to headphone position. For example, if a listener exhibited a bias which required a delay δTb to bring the image into the medial plane, then reversing either the headphones or the channels would require δTb to be introduced into the opposite channel for balance. It should be noted that the effect was barely noticeable owing to the large random variations in m between occasions. This was an interesting result in that it indicated no marked effect due to differences between headphones, although comparatively large physical differences had been observed in calibrating the instruments.

TABLE 5.11 : TEST FOR BIAS IN APPARATUS

	Delay Line in Left-Hand Channel	Delay Line in Right-Hand Channel
Total No. of Presentations	620	620
Judgments: "Left"	273	287
Proportion "Left"	0.440	0.463

Difference is not significant.

TABLE 5.12 : EFFECT ON m OF REVERSING CHANNELS AND HEADPHONES

✓ = "Correct" Connection ; X = Connections reversed

Subject	Headphones ✓		Headphones X	
	Channels ✓	Channels X	Channels ✓	Channels X
L.M.	-2, 1, -3, -2	-1, 1, 5, 7, 0, 1, 0, -2	0, -2 -1, -6	3, 6, 1, -1, 3, 5, 0, 6, 0, 2
T.M.	2, 0, -3, -5, -5, -12, 2, 5	3, 7, -1, 5 1, 3	0, 2, -1, -4, -2, 0 5, 8	-7, -10, 1, -2, -1, -10
K.P.	10, 11, 2, 6	-3, -8, -3, -8, -2, -7, 2, -3	-10, -13 -2, -7	5, 12, 5, 9, 3, 8, 5, 13 -6, -10
K.R.	5, 9, 10, 16 2, 5, 7, 9	-7, -13, -5 -9, 1, 4	-7, -9, 1, -6, -5, -7	9, 13, 5, 12
R.W.	-4, -6, -4, -5	5,5,5,2,5, 7,1,4,5,6 6,10,3,5,5	4, 5, 0, 2 1, 1, 2, 6, 4, 6	1,-1,-4,-4,-3 -5,-2,-4,-2,-4 -7,-9,-4,-5

Note: The values of m in the above table have been rounded off to the nearest unit. 1 unit = 5.8 microseconds.

Other checks for possible malfunctioning of the equipment were similarly carried out whenever convenient. These were generally favourable.

5.5 Stability of Performance during a Test

Testing for a change in regime is a difficult problem if one has no knowledge of how, when, or how many changes could have occurred, (see for example Quandt 1958,1960, Page 1955,1957). The method of subdividing the data into consecutive groups is useful in that it allows some check on regular trends in the data although the arbitrary grouping may blur the starting point of a trend and too many subdivisions of the data may lead to an inordinately high error variance. Changes in regime near the beginning or end of a set of data are also less likely to be noticed. As described, the previous experiments were analysed for time trends by the method of subgrouping the data. The results showed no significant change in b within an occasion and considering that in addition no significant change in b between occasions was observed it seemed reasonable to conclude that there was no marked variation in a listener's lateralization ability either during a test or between tests. An analysis of m, however, showed a large variation between occasions and possibly within an occasion. Although time-ordering did not indicate a highly significant effect in the overall analysis of variance, some subjects on certain occasions did show signs of a gradual drift in medial plane position.

A difficulty with the previous analysis was that in experiment A1 where 50 presentations per group were used, it was difficult to detect a trend in only 4 groups. In experiment A2 where 8 groups were used, only 25 presentations per group were available and the variability of the data was greatly increased. A similar analysis, but one which is more sensitive to regular trends is that of taking a "moving average" over the data. Successive groups of 50 consecutive presentations were analysed where the relative position of the group in the time sequence of the data was gradually shifted in steps of 25 presentations at a time. (The method was similar to that of carrying out a running autocorrelation except that probit analysis rather than simple correlation was used.) All the tests in experiments A1 and A2 involving a single sound source were analysed in this way. It was found that the data could be subdivided into two categories: those which showed no significant trend in m, as in Fig.5.2, and those which showed a marked trend, as in Fig.5.1. The b estimates showed a noticeable trend in only one case, which could reasonably have occurred by chance. The data in m suggested that whereas the value

FIG. 5.1 : EVIDENCE OF A TREND IN m

\- Test A26, Subject G.S. 1 Noise Source

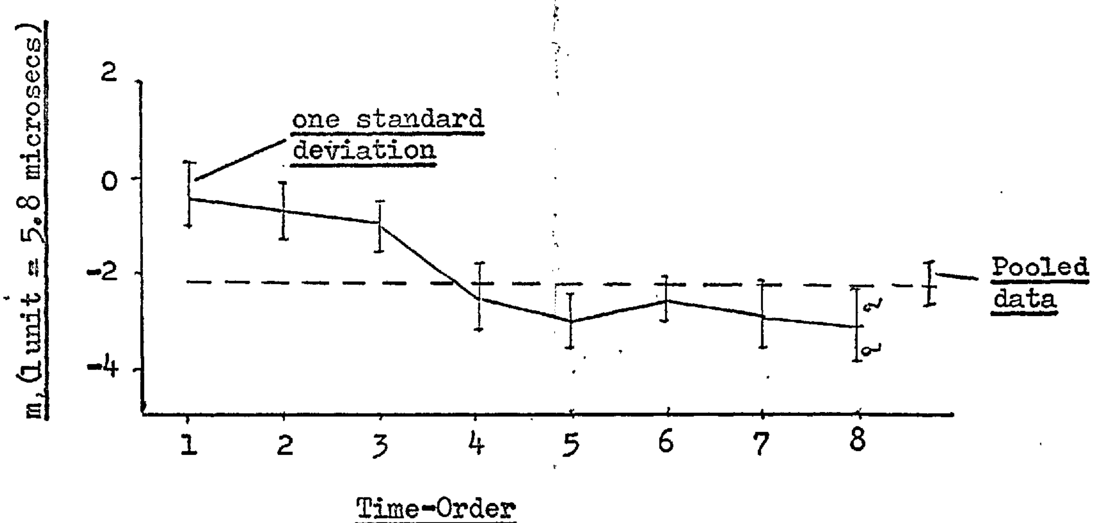
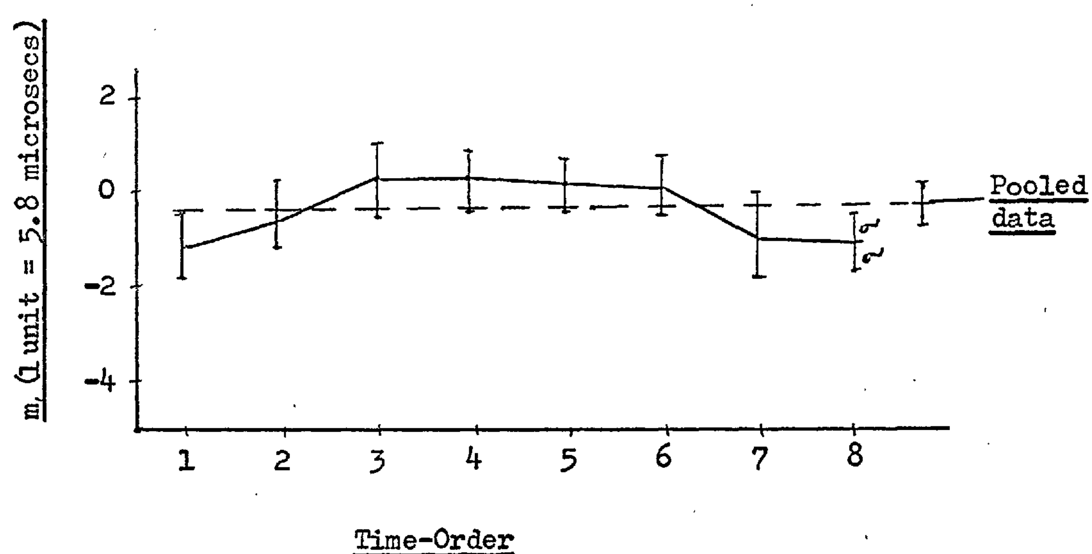


FIG. 5.2 : NO TREND IN m

\- Test A19, Subject G.S. 1 Noise Source



of δT corresponding to the position of the listener's medial plane was reasonably stable for most tests, on certain occasions (about 1 in 4) there was a gradual change in its value, i.e. there was a gradual drift in the position of the medial plane during the test. This result, however, could not be demonstrated conclusively owing to an inherent paucity of data. In subsequent experiments where highly efficient sequential strategies were employed, more convincing evidence of a drift in medial plane position was obtained.

5.6 Towards Optimum Data Allocation

A great advantage of subdividing the data is that it provides more independent observations and hence a larger number of degrees of freedom; furthermore, an additional factor, time-order, can be analysed. On the other hand, decreasing the number of data used in fitting the lateralization curve increases the sampling variance and also introduces a slight bias, (section 4.2.6). The most useful compromise depends on the inherent variability of the data which can only be determined empirically. The results of the previous experiments provided some useful information in this respect.

The effect of subdividing the data on estimation bias and error variance have been plotted in Figs.5.3 and 5.4 for the data of experiments A1 and A2. Clearly,

FIG. 5.3 : BIAS INTRODUCED BY DECREASING PRESENTATIONS/BATCH

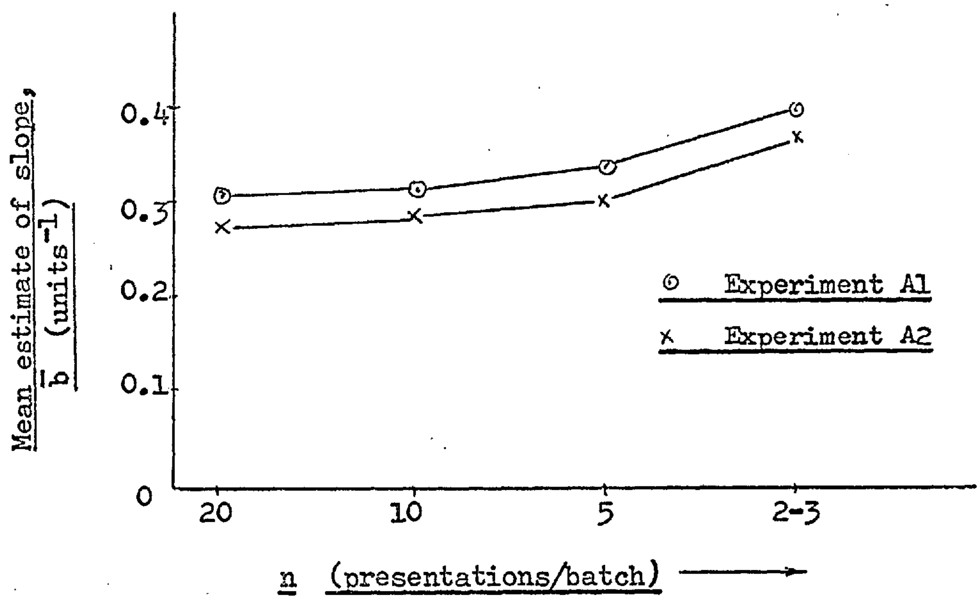
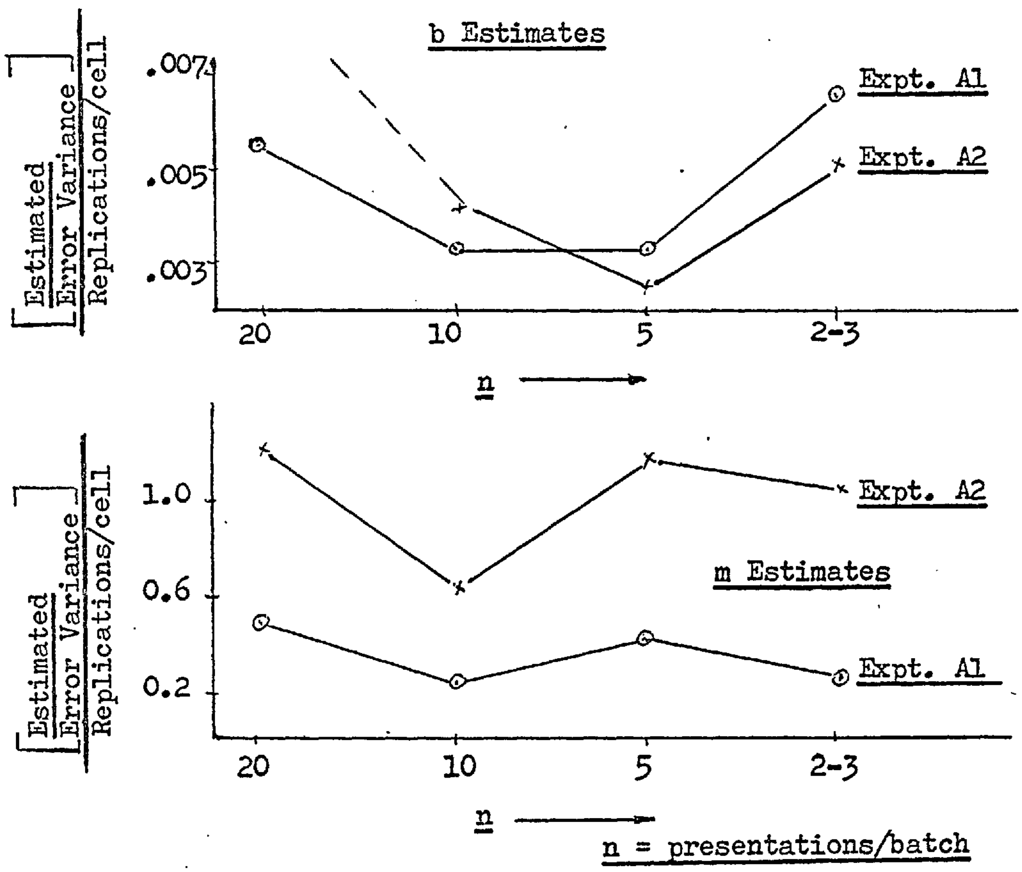


FIG. 5.4 : VARIATION IN EFFICIENCY



as the average number of presentations/batch, n, was reduced the overall error variance increased. For large n, sampling error played a comparatively minor role and its effect was swamped by other sources of error peculiar to the given experiment; hence the value of (estimated error variance)/(replications per cell) tended to be high. (Note that the number of replications per cell was inversely proportional to **n**). For smaller n, sampling errors obviously dominated and since sampling variance is roughly proportional to $\frac{1}{n}$, the curve of (estimated error variance)/(replications per cell) vs n tended to a minimum. For very small n, however, (< 5 presentations/batch) the linear relationship between sampling variance and n broke down; presumably because zero or 100% responses were so frequent. There was a sharp increase in estimated error variance at these values of n.

Assuming a fixed-effects model for the analysis of variance, the F-test reduces to a comparison between $\sigma_\varepsilon^2 + Mk$ and σ_ε^2, where σ_ε^2 is the error variance, k is the effect of the hypothesised factor and M the associated number of levels. M is proportional to the number of replications per cell; hence the F-test is most sensitive when (error variance)/(replications per cell) is a minimum. A similar result would be obtained if a component-of-variance rather than fixed-effects model were to be assumed. From the experimental data (Fig. 5.4) the best value of n for comparing estimates of slope appeared to lie in the range from 5 to 10 presentations/batch.

Another factor of importance was the bias introduced by using small samples. In Fig.5.3 the mean value of b averaged over all factors in the experiment has been plotted against n. For large n the bias was negligibly small, but for small n it was appreciable; in particular, there was a sharp increase for n less than 5. For n of 5 or 10, as recommended above, the bias was 5 - 10% which was not excessive.

5.7 Checking Statistical Assumptions

It was assumed that the lateralization curve was cumulative normal in the cross-over region. The validity of this assumption was checked using the data of the previous experiments. For each set of data the appropriate chi-square value was computed and the observed frequency distribution plotted. The results for experiments A1 and A2 are shown in Fig.5.5. Close agreement was found between observed and expected frequencies; strong evidence indeed that any deviations from the assumed distribution could be attributed solely to random experimental errors. The assumption of a

FIG. 5.5 : FREQUENCY DISTRIBUTION OF OBSERVED CHI-SQUARE VALUES

EXPT. A1 and A2.

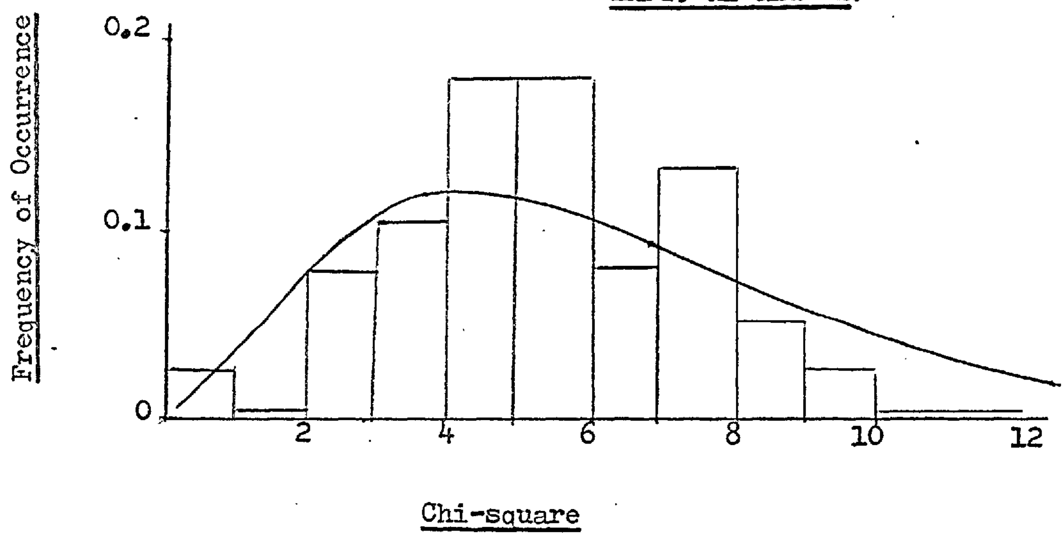
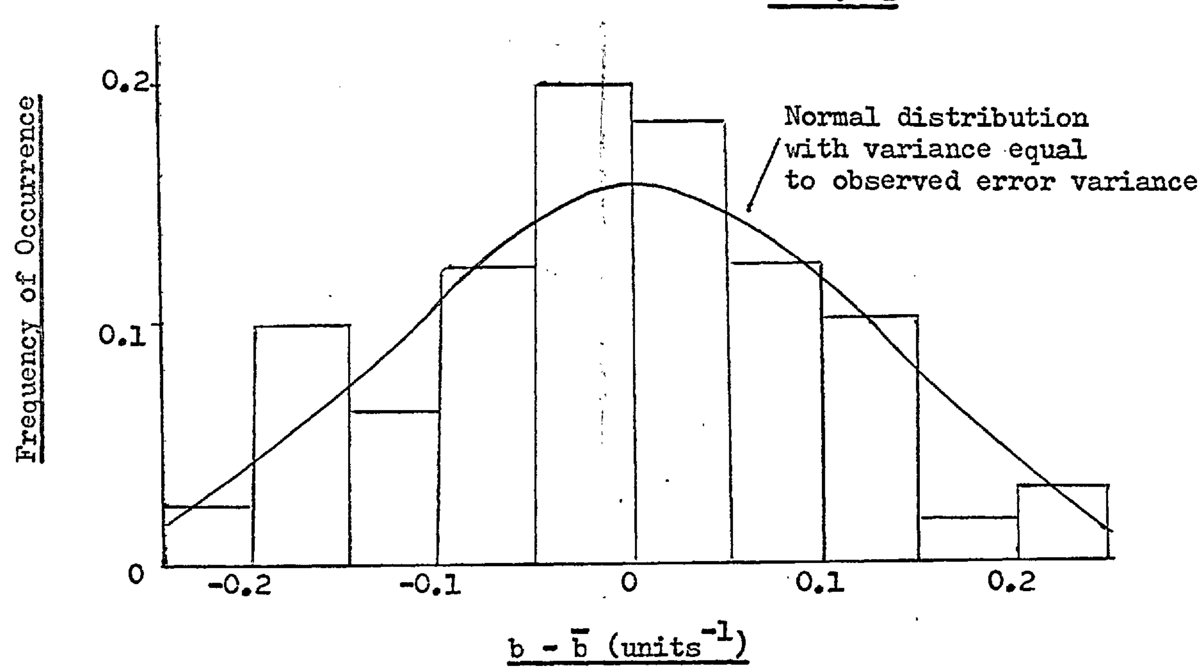


FIG. 5.6 : FREQUENCY DISTRIBUTION OF b ABOUT $\bar{b}$

EXPT. A1



cumulative normal distribution was therefore considered to be acceptable; other assumptions might also have been satisfactory.

In carrying out the analysis of variance it was assumed that the various estimates were normally distributed with equal variances. Having determined which factors were significant it was possible to plot the frequency distribution of the observed parameters about their estimated mean values. The results for experiment A1 are shown in Fig.5.6. A chi-square goodness-of-fit test indicated that the assumption of normality was reasonable. The assumption of homoscedasticity (equal variances) was more difficult to check. Theoretical considerations, however, (section 4.2.5) predict this to be the case for b (but not m) in an experiment involving batches of equal size and equi-spaced along the δT-axis.

5.8 Experiment A4. Images away from the Medial Plane

5.8.1 Purpose

Considering the apparent instability of the medial plane the possible use of other reference positions for the lateralization task was considered. The method of using a second, discriminable sound image as reference was employed; this enabled the experimenter to select the reference position at will. It was of particular interest to use this technique in order to measure relative lateralization ability for one of two sound images at some distance from the medial plane. Whispered speech was selected as the basic sound source for this investigation since the associated images were expected to be discriminable and at the same time not very different, physically, from random noise. Furthermore, should a subsequent investigation be carried out on the separability of successive approximations to normal speech (starting from modulated noise) then these results would be of value.

5.8.2 Description

Continuous, pre-recorded, male and female whispers were presented simultaneously. The listener was required to lateralize one about the other, having been told which image to use as the reference prior to each occasion. The positions of the two sound images for each presentation were selected at random subject to the constraint that the average value of δT for both images was fixed. The average value of δT was defined as

$\delta T_{av} = \frac{1}{2}(\delta T_1 + \delta T_2)$ where δT_1 and δT_2 were the interaural time delays for each of the two sound sources respectively. Since both δT_1 and δT_2 were selected at random the relative position of either image with respect to the medial plane could not in any way provide the listener with additional cues for the lateralization task.

The allocation of data followed the recommendations of sections 4.2.5 and 5.6. Five rather than ten presentations per batch were used since this was the smallest acceptable number and since it was expected that at least one factor would be insignificant in both b amd m (symmetry of equipment), in which case the resulting data could be combined to provide the more efficient value of 10 presentations/batch.

In this particular experiment, unlike the others, sound intensities were set by ear. In subsequent tests involving speech sounds, levels were set at 55 dB above threshold using a true r.m.s. meter with an averaging time of well over a minute. At the time of this experiment the instrument had not yet arrived.

The following 5 factors were investigated:

Average value of δT (3 levels; δT_{av} = 5, 35 and 75 units).

Sidedness (2 levels; both images either on the left-hand or right-hand side).

Subjects (3 levels; subjects G.S., P.N. and J.E.).

Reference Image (2 levels; reference either male whisper or female whisper).

Symmetry of Equipment (2 levels; channel 1 or channel 2 used for reference source).

The experiment was replicated 3 times.

Since an occasion was limited to 150 presentations, it was not possible to include all permutations of the factors within a single occasion. The less important factors (subjects, reference image and symmetry of equipment) were therefore confounded with differences between occasions. As it turned out, only the variation between subjects was found to be significant and hence this artifice was justified.

5.8.3 Results

The average value of δT and differences between subjects appeared to be the only factors which influenced b; m appeared to be independent of all factors, (Table 5.13). The data were combined for those factors which were insignificant in both m and b; the results are plotted in Fig.5.7. In order to provide additional points for plotting a curve, several ancillary observations were taken at $\delta T_{av} = 0$.

5.8.4 Discussion

There was a marked decrease in b with δT_{av}. Whether this was due to a "widening" of the sound image or to a smaller relative change in image position with

FIG. 5.7 : VARIATION OF b WITH δT_{av}

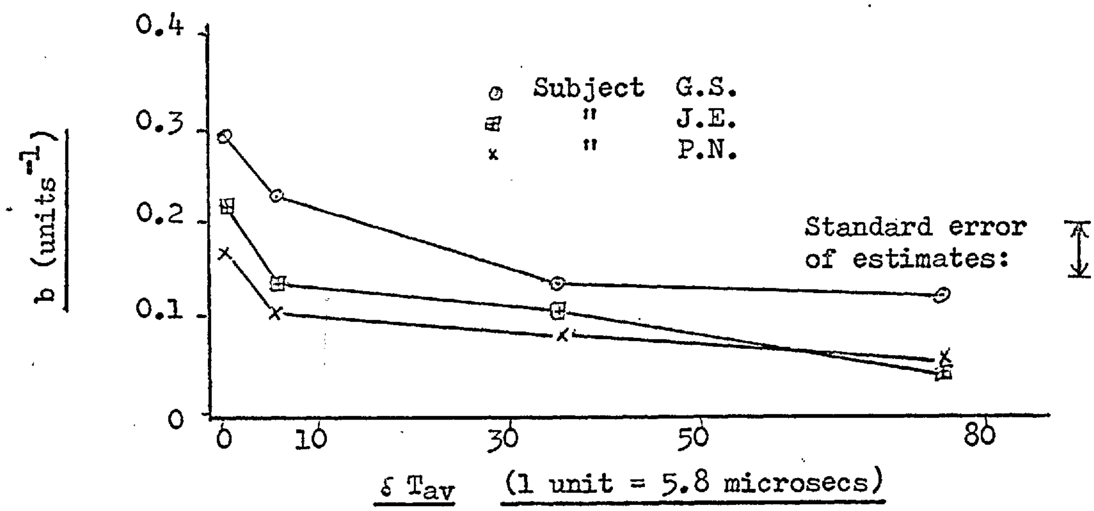


TABLE 5.13 : ANALYSIS OF VARIANCE, EXPT. A4

Effect	Variation in b		Variation in m	
	Mean Square	Significance Level	Mean Square	Significance Level
δT_{av}	.088	.995	13.2	-
Subjects	.327	.995	172.0	.995
Sidedness	.009	-	10.5	-
Reference Image	.0016	-	43.2	-
Symmetry of Equipment	.0102	-	1.0	-
Error Sum of Squares	.0106	-	21.1	-

Note: Interaction effects were insignificant.

δT, is difficult to say. In terms of δT, however, placing an image at some distance from the medial plane leads to a loss in image resolution. The listener appeared to be most sensitive to changes in interaural delay at $\delta T_{av} = 0$, where performance compared favourably with that for a single noise image lateralized about the medial plane. As before, differences between subjects were appreciable but of almost constant magnitude; this was another illustration of the apparent independence of those factors which influence lateralization ability.

The observation that sidedness had no significant effect on m or b was not unexpected. It indicated that the spatial location and resolution of sound images was symmetrical about the medial plane. Similarly it was not surprising to find no change in b whether a male or female whisper was used as the reference. This result was of interest since it showed that the factor confounded with it, occasions, was also insignificant in both m and b. This inference was further supported by the lack of any observed effect for the second factor confounded with occasions, namely, symmetry of equipment. The insignificance of the latter factor reflected favourably on the matching of the two sets of delay lines since a small regular difference in image position would have been readily observed.

The absence of any significant effect in m was noteworthy but not totally unexpected since the reference position had been fixed by the experimenter. In view of the greater stability obtained when using a second image as reference, it might be asked whether this artifice could not be used in measuring the degree of separability. The answer is a heavily qualified yes. The fact that a listener can provide meaningful lateralization judgements (i.e. $b > 0$) for a situation where no cues on lateral position are provided other than by the images themselves, indicates that the sound images are both separable and discriminable. The measurement of degree of separability, however, involves a comparison between the estimates of lateralization ability with one and with two sources, and it is not possible to avoid using the medial plane as a reference, with its attendant difficulties, in at least one of these measurements. This particular drawback could be avoided by using a specially selected reference image for all comparisons, but then other difficulties would be introduced; the reference image would be bound to interact with the sound images under test. Furthermore, the estimated ability of the listener to lateralize one sound image about another decreased markedly with average distance from the medial plane. Fig.5.7 shows that the effect was at least as large as the observed variation between subjects. The method of placing both images away from the medial plane would therefore introduce an additional nuisance parameter namely, distance from the medial plane.

In view of the above it was felt that the most appropriate avenue of research would be to seek out a method of tracking and compensating for gradual drifts in m during the course of a test. In this respect the use of a sequential strategy seemed promising.

The statistical properties of the data were investigated and found to be almost identical in form to that observed in the earlier experiments. The assumption of a cumulative normal was found to be justified. The most efficient allocation of data was estimated to be approximately 10 presentations/batch and the error variances of b amd m were both stable and nearly normally distributed. The results suggested that the statistical properties of the data may well be the same for a variety of different sound sources.

5.9 Summary and Conclusions

The primary aim of the preliminary experiments was to establish the repeatability of the technique and to determine the statistical properties of the data. In this respect the experiments were successful. The remarkable accuracy with which a listener canlateralize a sound image, whether whispered speech or random noise, was demonstrated. It was also estimated that there was a marked loss in lateralization ability if two noise sources were used in place of one, where one of the sources was presented throughout with a fixed value of δT such as to place the associated noise image near, if not in, the medial plane. The results were obviously repeatable and a high degree of precision was obtained.

Experimental factors which might influence lateralization ability were considered and their relative importance assessed. Differences between subjects were found to be appreciable but of constant magnitude. It was therefore considered advisable to carry out a comprehensive series of tests on a small crew of listeners rather than a few tests on many, randomly selected subjects. By using two sound sources which were fairly similar but obviously discriminable (male and female whispered speech) it was also shown that lateralization ability appeared to decrease with distance from the medial plane. Other factors such as ordering effects within a test or differences between tests were found to have an insignificant effect on b, although this was not true for m. It was noted that those factors which did influence lateralization ability appeared to act independently.

The statistical properties of the data were analysed in some detail. The assumption of a cumulative normal was found to be justified, and checks on other statistical assumptions were also favourable. An efficient method of allocating the data was derived; it was used in experiment A4 and presumably would be of value in subsequent investigations.

Checks were carried out for possible effects due to a lack of balance or other malfunction of the equipment. The results were favourable. Although careful experimental design can balance out the effects of consistent biases in the equipment, factors which vary at random between presentations are implicitly confounded with the main effect being measured, i.e. lateralization ability. Stability of the equipment was therefore of great importance.

Possible instability of the position of the medial plane was a matter of some concern since it would have biased estimates of lateralization ability. Furthermore, there was considerable doubt as to whether the fixed sound image could be reliably placed in the medial plane. Although the experimental results did not show conclusively that there had been a gradual shift in medial plane position on certain occasions, the effect appeared to be not unlikely. The ambiguity of this result was due largely to an inherent paucity of data, which in turn was due to the limited duration of a test. If a second sound were to be used as the reference in place of the medial plane, then greater stability would ensue, but at the cost of additional difficulties. A direct comparison with the estimated lateralization ability for a single source would not be possible and if a special reference image were to be introduced, some form of interaction would be inevitable. It was felt that the most promising approach would be to seek out a technique of tracking and compensating for changes in medial plane position during the course of a test. A sequential strategy seemed an appropriate technique and was considered subsequently.

In brief, it appeared that the proposed measure of separability had many practical advantages, (accuracy and repeatability of measurements, consistency of assumptions), but that the possible instability of the medial plane and the associated problem of placing a sound image sufficiently close to it were major difficulties which had to be resolved.

CHAPTER 6 NOISE IMAGES AND THE CORRELATION PROCESS

6.1 Introduction

An experimental investigation often provokes more questions than it purports to answer. Of the many questions raised in the previous investigation the following was rather provocative. The difference limen along the δT axis for a single noise source has been estimated to be approximately 3 units. When two random noise sources were used differences between the associated interaural delays were as much as 5 units yet no subject claimed to hear more than one image. It has been shown in numerous investigations that if two physically similar sound sources are used a single image is heard, even for very large differences in the associated interaural delays; e.g. the Precedence Effect. Could it be that this effect also occurs with sources which are similar in average terms only, such as two statistically identical noise sources? Theoretically it is possible by means of say a running crosscorrelation between the signals at the two ears, to detect whether one or two sources are being used - provided there is a difference in the associated interaural delay for each source.
It was of interest to determine whether a human listener could detect a difference between one or two noise

sources and, if so, whether in fact two separate images were perceived.

Several other points were of interest, not least of which was the problem of finding a method for tracking shifts in medial plane position during the course of a test. The possible interpretation of the results in terms of a correlation model of the Cherry-Sayers type was also considered. The purpose of this series of experiments (Series B) was therefore threefold:

a) to investigate whether two separate but statistically identical noise sources produce separate noise images,

b) to interpret the results in terms of a simple correlation model, and

c) to seek out a method for tracking gradual shifts in the position of the medial plane during the course of a test.

6.2 Experiment B1, Ability to Detect whether One or Two Sources are Presented

6.2.1 Purpose

To determine the ability of a listener to detect whether one or two statistically identical noises are being presented, given a difference between the interaural delays associated with each source.

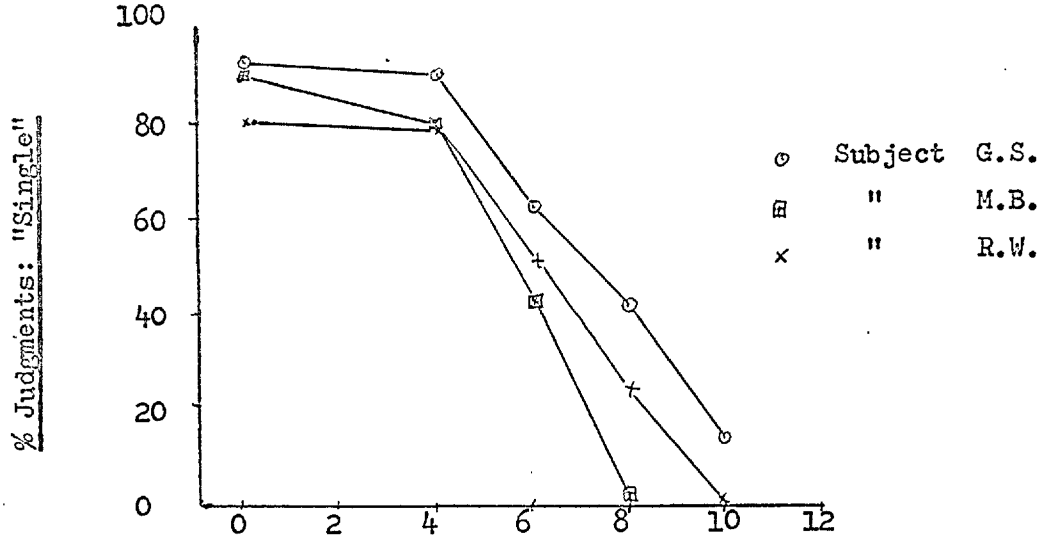
6.2.2 Description

Separate tape recordings of random, white noise subsequently bandlimited from 125 to 4000 c.p.s. at a signal level 50 dB above threshold were used for each independent source. Each presentation consisted of the two sources being switched on simultaneously, remaining on continuously until the listener made the simple binary judgement, "single : dual source", after which the stimuli were switched off simultaneously. The interaural delays for each sound source, δT_1 and δT_2, were varied at random between presentations subject to the constraint that $\delta T_1 + \delta T_2 = 0$. A total of 40 presentations/batch were used. ($\delta T_1 = \delta T_2 = 0$ was an exception with 32 presentations). Batch positions of 0, $\pm$ ~~4~~ 2, $\pm$ ~~6~~ 3, $\pm$ ~~8~~ 4 and $\pm$ ~~10~~ 5 units were decided on after a pilot experiment on one of the subjects (G.S.). 3 subjects were used and each was subjected to a trial run prior to the test.

6.2.3 Results and Discussion

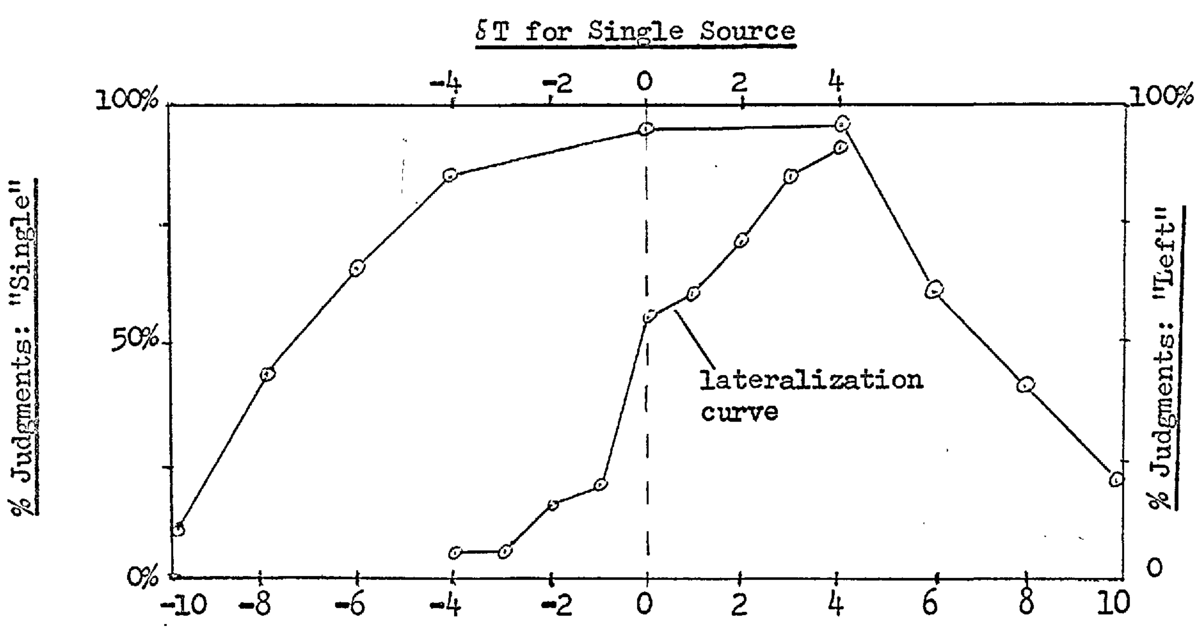
The proportion of decisions indicating a single source as related to the distance between δT_1 and δT_2 is plotted in Fig.6.1. It should be noted that with $\delta T_1 = \delta T_2$ the two sources are theoretically indistinguishable from a single source of twice the power. Although the proportion of "single source" decisions in the vicinity of $\delta T_1 - \delta T_2 = 0$ was very high, it never

FIG. 6.1 : DISCRIMINATION BETWEEN ONE AND TWO SOURCES - EXPT. B1



Distance between sources ($\delta T_1 - \delta T_2$)
1 unit = 5.8 microseconds

FIG. 6.2 : LATERALIZATION ABILITY AND ABILITY TO DISTINGUISH ONE FROM TWO SOURCES



Distance between sources ($\delta T_1 - \delta T_2$)
1 unit = 5.8 microsec.

quite reached unity. Several spot checks were carried out during the experiment and it was found that the subjects quite often gave a "dual source" decision when only one source was in fact used. These decisions were referred to as false positives and have been considered in greater detail in experiment B3.

While carrying out the pilot tests it was noted that if a very wide spacing in $\delta T_1 - \delta T_2$ was employed, no false positives were observed. It appeared that the frequency of false positives depended in part on the average difficulty of the listening task. In this and subsequent experiments the range of $\delta T_1 - \delta T_2$ was selected so as to just exceed the transition region, thus making the listening task consistently difficult. Under these conditions the listeners were perhaps more sensitive to "dual source" decisions; hence the occurrence of false positives. Since the aim of the experiment was to measure the upper limit of sensitivity to changes in $\delta T_1 - \delta T_2$, this was felt to be not unsatisfactory.

A most noticeable feature of the results was that although the listener was very sensitive to changes in $\delta T_1 - \delta T_2$, the range over which two sources were indistinguishable from one was just wider than the transition region of a typical lateralization curve, (See Fig.6.2). This would explain why in Series A when two noise sources were used the listener did not claim to hear more than one noise image.

The knowledge that no more than two noise stimuli were being applied was of great assistance to the listener since any noticeable difference in the quality of the sounds could be used as a cue. The subjective impressions of the listeners, as recounted immediately after the tests, suggested that the width of the sound image was commonly used as the criterion of whether one or two noise stimuli were presented. It was obviously of interest to determine whether separate images could be heard for sufficiently large differences in δT. This formed the basis of the subsequent experiment.

6.3 Experiment B2. The Separability of Multiple Noise Images

6.3.1 Purpose

To determine whether two or more random noise sources which are statistically identical can produce separate sound images occupying distinct regions in subjective space.

6.3.2 Description

Three pre-recorded, low-pass, random noise sources ($F_c \doteqdot$ 4000 c.p.s.) were used, each at a level of 50 dB above threshold. Two of these were presented with values of δT of +100 and -100 units respectively; a negative value indicating a delay in the left-hand channel. The interaural delay for the third source was varied at random over the range -10 to +10 units. 5 or 6 batches at 10 presentations per batch were used and a total of 5 subjects tested.

The listener was required to lateralize the central image about the medial plane. A short training session preceded each test in which the third source ($-10 < \delta T <$ $+10$ units) was switched in and out at random (between presentations); the subject was required to state whether or not the third source was present. When it was clear that the listener could accurately detect the presence of the third source the test proper was begun.

6.3.3 Results and Discussion

Three sources were used since it enabled identification of the images. In this way it was possible to measure the degree of separability of one of the three images. Had only two sources been used the subject might have been able to use a criterion other than image position for the lateralization judgements, e.g. if the two sources yielded a single image then the listener could have used the mean position or the width of the image as a cue, with similar results.

The data indicated (Table 6.1) that the subject could clearly track the position of the third image, although there was a marked decrement in the estimated degree of separability. Instability of the medial plane presumably played some part in decreasing the resolution of the measurements, but in this case the interaction between the wanted and unwanted images was so large as to swamp its effect. It was clear that the observed results could not have been obtained had a single, diffuse sound image been perceived. Hence it was concluded that separate but statistically identical noise sources could produce separate sound images - provided there was a sufficiently large difference in spatial position.

TABLE 6.1. Separability of Central Noise Image

Subjects	Estimated Lateralization Ability		Estimated Degree of Separability $b_A - b_{A/B}$	$\frac{b_A - b_{A/B}}{b_A}$ x 100
	Single Source b_A	1 of 3 Sources $b_{A/B}$		
G.S.	.337	.176	.161	47.7%
M.B.	.335	.098	.237	70.7%
R.W.	.811	.182	.629	77.6%
L.M.	.678	.126	.559	81.4%
K.R.	.659	.089	.570	86.5%

values of b are quoted in units of 5.8^{-1} microseconds^{-1}

6.3.4 Ancillary Tests

Informal tests were carried out which indicated a similar effect using interaural intensity rather than time differences. Another short test showed that a subject could reliably judge whether 1, 2 or 3 sources were being used. The relative positions of all three were varied at random subject to the constraint that difference between the interaural delays were always greater than 50 units.

6.4 Interpretation of Results in terms of a Correlation Model

6.4.1 Introduction

Correlation models have been prominent in interpreting and predicting the results of binaural experiments (e.g. Cherry and Sayers 1957, Voelcker 1961, Licklider 1951). The model proposed by Cherry and Sayers is particularly valuable since it is simple in concept, covers a wide range of experimental data and is capable of extension. Briefly, the model postulates a running autocorrelation at each of the two inputs and a running crosscorrelation between the outputs of these autocorrelators. A simple decision rule based on the relative proportion of the correlation pattern to the left or right of the $\tau = 0$ axis predicts the frequency of left or right judgements by the listener. The effect of interaural intensity differences is accounted for empirically by truncating the pattern in the vertical axis; this is perhaps the most arbitrary aspect of the model. The nature of the time window (i.e. time weighting function) used in the running correlation is not fully specified and this allows for some flexibility in fitting the model to observed data.

The results of experiments B1 and B2, however, suggest a model involving a direct crosscorrelation between the two inputs. If autocorrelation precedes crosscorrelation then the differences between one and two statistically identical noise sources would be lost in the formation of a single peak in the autocorrelation pattern for each input. This result does not preclude the possibility of a running autocorrelation but rather that the two correlation processes are more likely to operate in parallel than in series. Furthermore, those situations containing important cues which can be isolated by means of one or other correlation process may be interpreted separately. For example, many monaural effects could be predicted by a model involving autocorrelation only; on the other hand the data on the separability of noise images may be interpreted solely in terms of a running crosscorrelation. Such a model is here postulated.

6.4.2 Description of the Model

A running crosscorrelation between the two inputs is hypothesized. The resulting correlation function is interpreted in the following way:

a) The subjective impression of a sound image corresponds to a peak or maximum in the correlation function, although the converse need not necessarily be the case (e.g. the multiple peaks for a low frequency tone).

b) The vertical axis at $\tau = 0$ corresponds to the expected position of the medial plane. The apparent fluctuation of the position of the medial plane about the mean value is due to a variety of factors, random or otherwise, including mounting of the headphones.

c) The position of a peak relative to the position of the medial plane, corresponds to the lateral position of the sound image.

d) The spread of a peak is defined as that distance along the τ-axis between the maximum and a point on the correlation function at a critical depth, c, below the maximum (see Fig.6.3). The spread corresponds to that subjective quality of a sound image referred to as width, a wider image occupying a greater region in subjective space.

e) The height of a peak, which is numerically equal to the short-term average power of the associated sound

FIG. 6.3 : TYPICAL CORRELATION PEAK

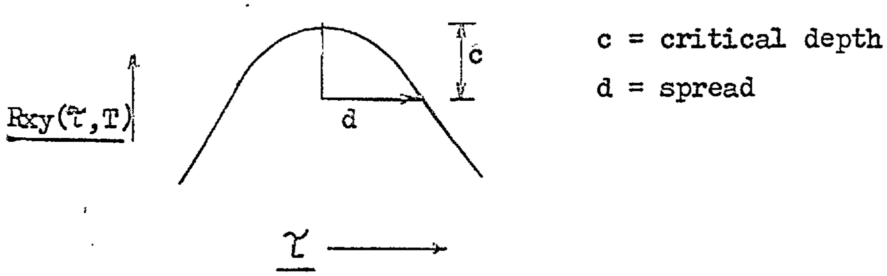
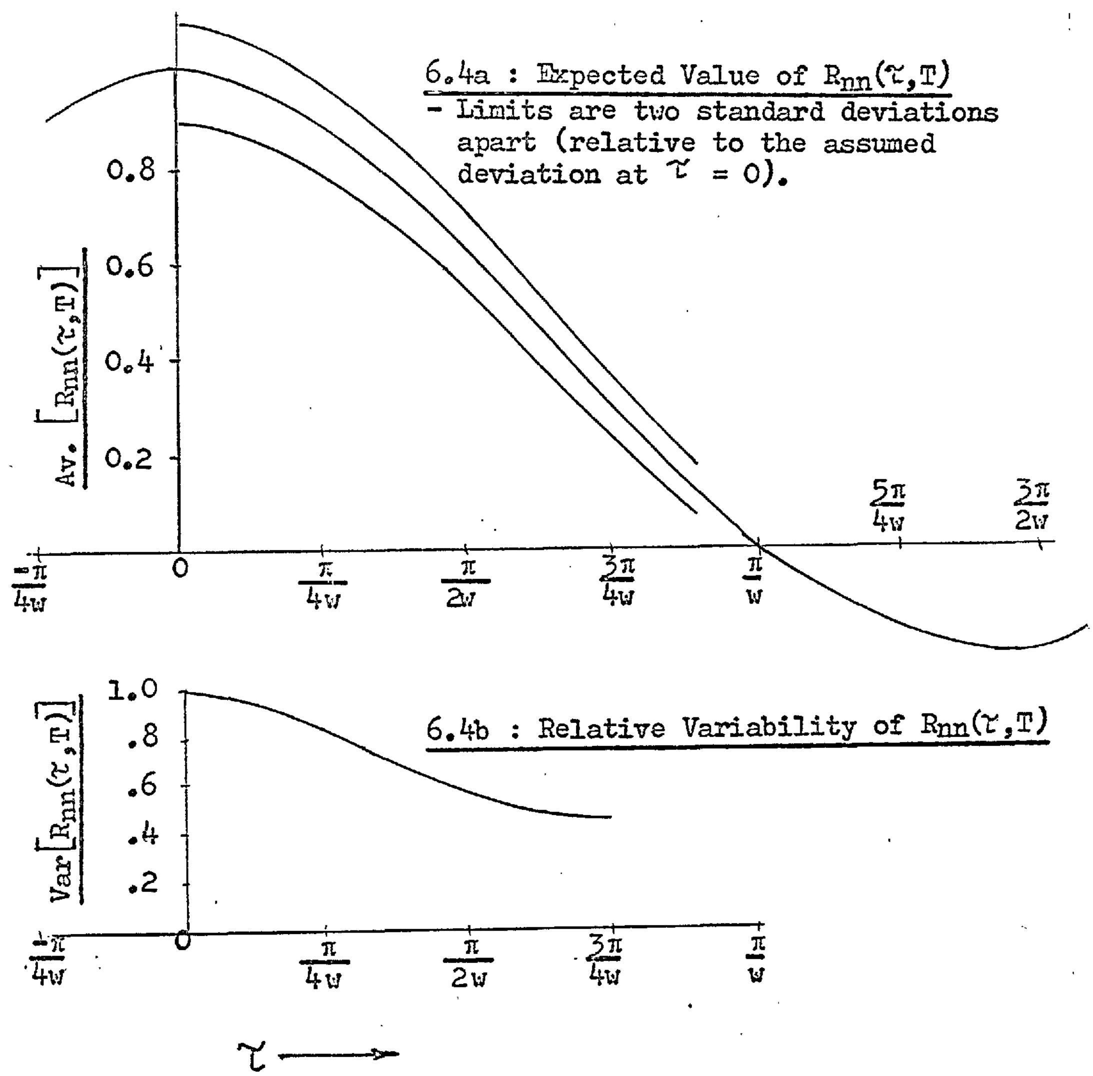


FIG. 6.4 : RUNNING AUTOCORRELATION OF LOW-PASS BAND-LIMITED NOISE



source, corresponds to the loudness of the sound image.

f) The variability of a listener's responses is due both to fluctuations in the correlation pattern resulting from the finite averaging time, and to a component of variability characteristic of the listener.

No specific time window for the running correlation has been postulated, but for simplicity a rectangular window may be assumed. The effective length of the time window, however, needs to be estimated from the data. Preliminary considerations suggest an averaging time of the order of milliseconds.

The postulated model is clearly an extension of that proposed by Cherry and Sayers, but there are important differences even where the models overlap. The predicted variability of a listener's judgements in the present model depends both on the variability of the running correlation function and on other unspecified sources of uncertainty characteristic of the listener. In the Cherry-Sayers model the expected frequency of a listener's lateral judgements depends solely on the shape of the correlation function.

The underlying aim of the proposed model has been to find suitable physical correlates for the subjective quality of separability and, in particular, to provide a satisfactory interpretation of the previous experiments. In this respect it has been successful.

Fig.6.4 shows the correlation function for a random, white noise subsequently bandlimited to $W/2\pi$ c.p.s. and with an associated interaural delay $\delta T = 0$. In calculating the relative variability of this function (Appendix A1) a variance of unity was assumed at $\tau = 0$, without loss of generality. The absolute value of the variance is inversely proportional to the bandwidth, W, and the effective averaging time, T. If two sources are used with a small difference in their associated interaural delays the two peaks in the associated correlation pattern may be so close together that the dip between them is blurred by the variability of the pattern. In this case only one peak would be apparent, but having a wider spread than either of the original peaks.

For three sound sources each with a markedly different value of δT, three distinct peaks occur in the correlation pattern. The variability of the pattern in the region of the central peak, however, is greater than that for a single source. This is due to the additional variability of the tail ends of the outlying peaks and also, in part, to the finite crosscorrelation between independent noise components. The greater variability of the central peak clearly predicts the poorer performance of the listener in lateralizing one of three sound images.

The model as postulated consists of several unknown quantities which can only be estimated from the data; critical height c, averaging time T, and a constant characteristic of each listener. The data of the previous experiments were too few to provide a quantitative check on the model and for this reason experiments B3 and B4 were devised. Bandwidth was chosen as the controlled variable since it was convenient to use and had a marked effect on the spread and variability of the correlation function.

6.5. Experiment B3. Effect of Bandwidth on Ability to Discriminate between One and Two Noise Sources

6.5.1 Purpose

To investigate the effect of bandwidth on the ability of a listener to detect whether one or two statistically identical noise sources are being presented and to relate the results to the proposed model.

6.5.2 Description

The method was essentially that of experiment B1. Low-pass, random noise was used with values of cut-off frequency Fc = 900, 1350, 2025, 3040 and 4560 c.p.s. respectively. Each occasion consisted of a single test, the bandwidth being fixed for each test. A control batch consisting of presentations from a single source of twice the power was included. Four subjects were tested, (originally five, but one, a schoolboy, gave totally incoherent results and his data was omitted in the final analysis). Of these subjects, two were particularly experienced and two were novices; an additinal factor - experience of subjects - was therefore analysed.

A standard factorial design was used, the factors being:

Bandwidth (5 levels; 900, 1350, 2025, 3040 and 4560 c.p.s.)

Subjects (2 levels; 1st and 2nd subject)

Experience of Subjects (2 levels; experienced vs inexperienced).

Several pilot tests were carried out in order to estimate a suitable range for the difference in interaural delays.

6.5.3 Results and Discussion

The results given in Tables 6.2 to 6.4 and Fig.6.5, are in accord with the predictions of the model; ability to distinguish one from two sources is ~~inversely~~ proportional to bandwidth. In fitting the data, there were 5 unknowns, (a constant for each subject and a constant characteristic of the correlation pattern) and 20 independent observations. This allowed 15 degrees of freedom for testing the goodness of fit. It was

TABLE 6.2 : DIFFERENCE LIMENS BETWEEN ONE AND TWO NOISE SOURCES - EXPT.B3

Bandwidth (c.p.s.) 900	1350	2025	3040	4560	Subjects	
12.1	7.3	7.4	5.3	5.2	R.W.	Experienced.
16.6	14.0	12.0	9.0	10.6	L.M.	Experienced.
9.3	8.8	10.8	8.5	9.0	I.C.	Inexperienced.
16.0	16.5	11.7	13.5	10.2	E.W.	Inexperienced.

Results of Pilot Study:- (Bandwidth - 4560 c.p.s.)

4560	Subject
5.4	R.W.
9.8	I.C.
7.9	E.W.

Note: The difference limen is the distance at which 50% of the judgments indicate a dual source. 1 unit = 5.8 microsecs.

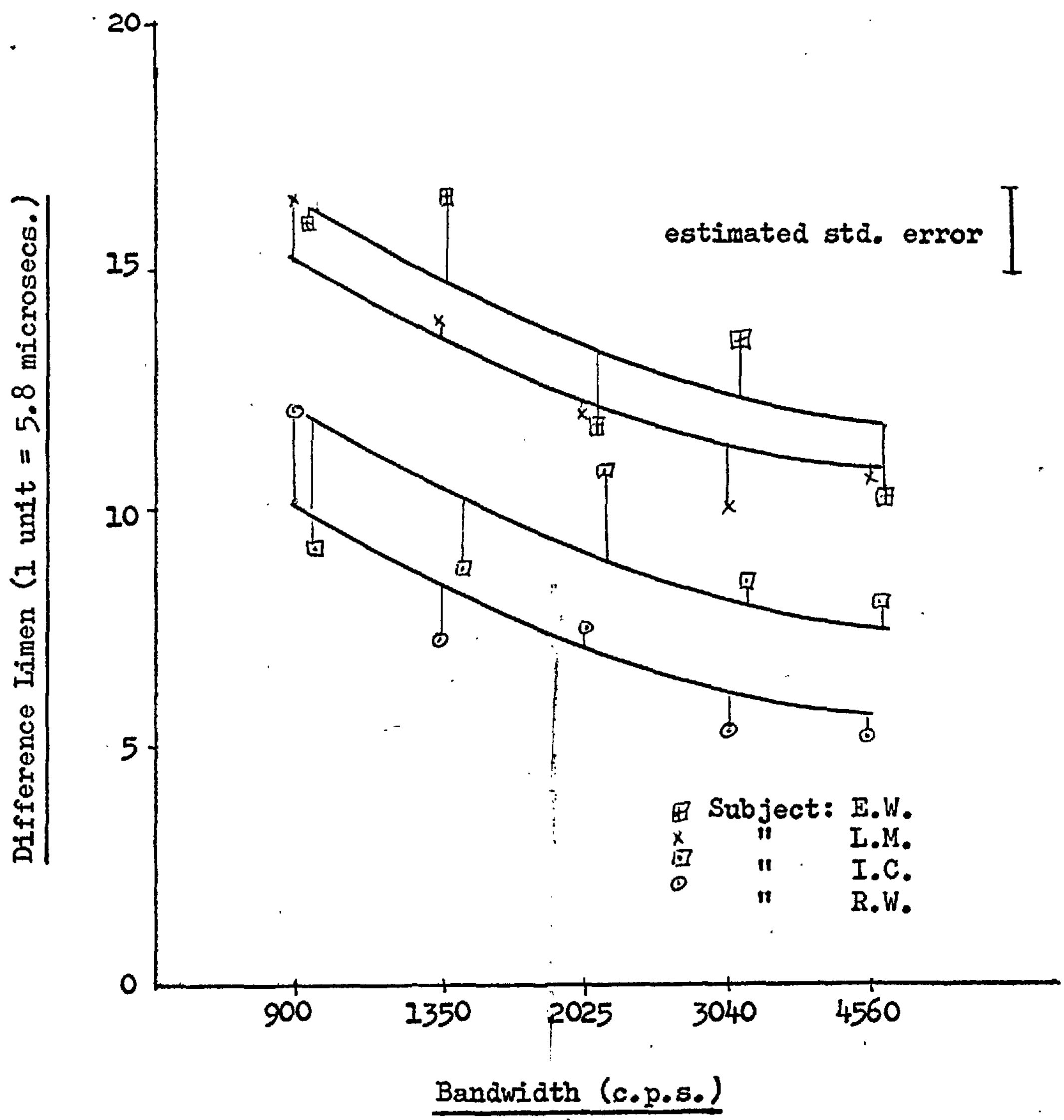
TABLE 6.3 : ANALYSIS OF VARIANCE - EXPT.B3

Effect		Mean Square	Significance Level
Bandwidth	(A)	14.3	.95
Subjects	(B)	111.6	.995
Experience	(C)	9.5	.9
Interactions:	A x B	3.2	-
	B x C	0.8	-
	A x C	3.5	-
Residual		2.1	

TABLE 6.4 : PROPORTION OF FALSE POSITIVES

Subject	Single Source	Dual Source $\delta T = 2$	Results of Pilot Study: Single Source	Results of Pilot Study: Dual Source $\delta T=0$
R.W.	.10	.30		
L.M.	.38	.32		
I.C.	.39	.46		
E.W.	.48	.50	.42	.47
Estimated Std.Error	.065		.064	

FIG. 6.5 : DISCRIMINATION BETWEEN ONE AND TWO NOISE SOURCES.



found that the mean square deviation of the observed data from predictions of the model could reasonably be attributed to sampling errors alone, (see Appendix A1).

Differences between subjects were, as usual, found to be highly significant; the experience of the subjects was also an important factor. This effect was further demonstrated in a lower proportion of false positives registered by the experienced subjects, (Table 6.4). In analysing the data on false positives no significant difference was found for two superimposed noise sources and one source of twice the power. It was also observed that the frequency of "dual source" decisions at a spacing of $\delta T_1 - \delta T_2 = 2$ units was not significantly different from that for a single source, and the data was combined to provide a more reliable estimate for the frequency of false positives. The technique suggested by Finney (1947) for compensating for the effect of false positives was used in estimating the j.n.d. (50% dual source decisions) in the transition from single to dual responses. The rate of false positives was, unfortunately, comparatively high thus increasing the variability of the estimates.

6.6 Experiment B4. Lateralization Ability and Noise Bandwidth

6.6.1 Purpose

To investigate the effect of bandwidth on the slope of the lateralization curve for low-pass noise, and to relate the results to the proposed model.

A second aim was to use an efficient sequential strategy in order to determine whether or not a drift in the midpoint of the curve occurred during a test.

6.6.2 Description

A sequential estimation technique was now employed for the first time. The strategy used was Mood and Dixon's simple Up-and-Down method; i.e. a reversal in the direction of changing δT follows a single change in response type. Steps of 1 unit were used. Groups of 4 runs were combined to provide reasonably independent m estimates. The estimated slope, b, was determined by probit analysis.

Low-pass noise of bandwidths 900, 1350, 2025, 3040 and 4560 c.p.s. were used, the tests being carried out on 5 subjects. As usual, a factorial design was employed, the factors being:

Time-Order (5 levels; 1st, 2nd .... 5th group of 4 runs)

Subjects	(5 levels; R.W., L.M., E.W., I.C., and S.L.)
Bandwidth	(5 levels; 900, 1350, 2025, 3040 and 4560 c.p.s.).

6.6.3 Results and Discussion

The results were rather surprising in that estimated lateralization ability (Tables 6.5 and 6.6) appeared to be quite independent of bandwidth. Considering that the spread of the correlation peak is inversely proportional to bandwidth a change in image resolution and hence lateralization ability was expected. In terms of the model, therefore, this result indicated that the component of variability characteristic of the listener was very much greater than those effects related to the correlation pattern - at least for low-pass noise with a cut-off frequency in the range 900 to 4560 c.p.s.

The sequential strategy was highly efficient in estimating the midpoint and the results showed a marked drift in the estimated position of the medial plane on certain occasions, (Tables 6.7 and 6.8). The significance level for the effect of time-order on m exceeded the 99% level; this is to be contrasted with the ambiguous result of Series A where a drift in m seemed likely but not probable.

TABLE 6.5 : LATERALIZATION ABILITY AND NOISE BANDWIDTH

Bandwidth (c.p.s.)					Subjects
900	1350	2025	3040	4560	
.358	.825	.357	.472	.335	R.W.
.752	.126	.159	.197	.531	L.M.
.108	.360	.183	.521	.417	I.C.
.400	.277	.380	.242	.294	E.W.
.148	.495	.397	.522	.198	S.L.

All values quoted in units of 5.8^{-1} microseconds^{-1}.

TABLE 6.6 : ANALYSIS OF VARIANCE - LATERALIZATION ABILITY

Effect	Mean Square	Significance Level
Bandwidth	.0105	No factor
Subjects	.0195	significant.
Residual	.0426	

TABLE 6.7 : UP-AND-DOWN ESTIMATES OF MIDPOINT FOR VARIOUS NOISE BANDWIDTHS

Time-Order 1	2	3	4	5	Bandwidth c.p.s.	Subjects
4.75	5.5	6.25	7.5	8.5	900	S.L.
-9.0	-9.25	-9.0	-7.5	-8.5	1350	
1.75	3.0	4.5	5.25	5.0	2025	
2.25	3.75	3.75	3.5	4.5	3040	
-1.75	-2.5	-2.0	1.0	1.25	4560	
1.5	2.5	3.0	4.75	8.0	900	I.C.
7.5	7.5	6.25	6.0	7.0	1350	
0.5	1.75	3.75	4.75	4.75	2025	
-0.25	0.25	1.0	0.5	-0.25	3040	
1.5	-0.25	-0.5	0.25	1.5	4560	
-1.25	-2.0	-2.0	-2.0	-0.75	900	R.W.
-4.0	-5.0	-5.0	-5.75	-5.75	1350	
-0.75	-0.75	-1.0	-2.25	-2.0	2025	
-1.25	-1.75	-2.75	-3.5	-4.0	3040	
0.5	2.75	4.0	3.75	4.0	4560	
-4.25	-4.25	-4.25	-4.0	-3.25	900	L.M.
-4.75	-5.5	-7.25	-9.5	-9.5	1350	
-5.0	-5.0	-4.75	-7.5	-9.25	2025	
7.75	8.0	9.5	9.5	9.75	3040	
-3.25	-3.5	-3.5	-4.5	-4.25	4560	
6.75	8.0	8.5	8.75	8.5	900	E.W.
-8.0	-6.25	-5.5	-4.25	-4.25	1350	
-0.25	2.0	2.5	1.5	2.25	2025	
-2.25	-0.75	1.0	0.5	0.25	3040	
-7.5	-8.5	-7.75	-6.0	-6.0	4560	

Note: Standard Error of above estimates $\doteqdot$ 1.5 units.

1 unit = 5.8 microsecs.

TABLE 6.8 : ANALYSIS OF VARIANCE IN m

Effect		Mean Square	Significance Level
Time-Order	(A)	8.36	.99
Bandwidth	(B)	361.4	.995
Subjects	(C)	228.9	.995
Interactions:-	A x B	3.66	just .9
	B x C	6.70	.95
	A x C	238.4	.995
Residual		2.26	

It was possible that the results obtained in this experiment were partly a function of the experimental technique. The Up-and-Down strategy is a particularly simple one and the subjects may have recognized a pattern and hence adjusted their responses accordingly. One subject commented quite spontaneously that the sound image appeared to have shifted gradually from one side of the head to the other and only reversed direction after a change in response; which indeed supported the above fears. Moreover, the estimation strategy was most efficient for estimating the midpoint and the sampling error of the slope estimates may have been unnecessarily large. For these reasons, it was decided to repeat part of the experiment using a variation of the newly developed Wetherill technique which is not subject to the above objections. Using this technique two points on the curve could be estimated thus allowing for an efficient estimate of σ (= $^1/b$) and at the same time concealing any evidence of a sequential strategy by interleaving presentations at random (see section 4.3.3). The measurements at the end of the range, i.e. Fc = 900 and 4560 c.p.s., were to be repeated.

6.7 Experiment B5

6.7.1 Purpose

To check the results of experiment B4 using an alternative sequential strategy.

6.7.2 Description

The Wetherill strategy, as described in section 4.3.2, was used. Steps of 1 unit were employed and the sequence was terminated after 13 runs, thus allowing 6 independent estimates. Three subjects were tested using low-pass noise of 900 and 4560 c.p.s. bandwidth respectively. The factors tested were:

Time-Order	(6 levels)
Bandwidth	(2 levels; 900 and 4560 c.p.s.)
Subjects	(3 levels; L.M., R.W. and K.R.).

6.7.3 Results and Discussion

The results wereessentially the same as those of the previous experiment. No significant change in s (= $^1/b$ = estimate of σ) with bandwidth was observed (Tables 6.9 and 6.10), although two very different values of bandwidth were used. A significant drift in m during a test was observed on one of the six occasions(Fig.6.6), and a possible drift on at least two other occasions. In the overall analysis of variance the stability of m in the majority of the tests swamped this effect and an insignificant time-order sum of squares was obtained.

TABLE 6.9 : EFFECT OF BANDWIDTH-WETHERILL ESTIMATES

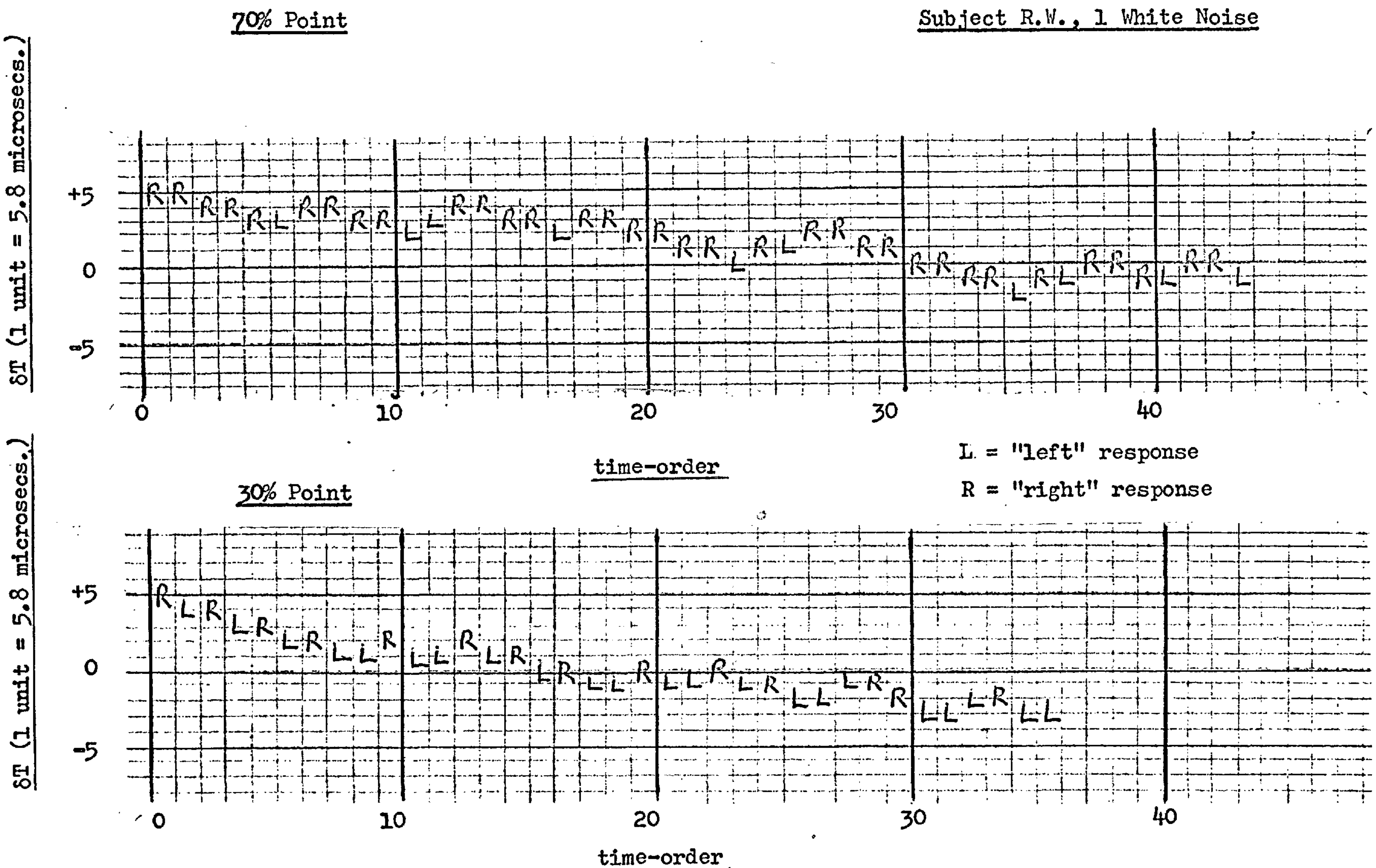
Time-Order 1	2	3	4	5	6	Bandwidth c.p.s.	Subjects
Estimates of midpoint							
4.25	3.75	4.0	4.5	4.75	4.5	900	R.W.
-4.75	-5.25	-6.5	-7.5	-8.25	-8.5	4560	
1.5	1.5	0.5	-1.0	-1.0	-1.5	900	L.M.
3.5	3.75	2.25	1.75	2.5	-0.75	4560	
12.0	12.0	12.0	13.0	14.0	15.0	900	K.R.
10.5	12.0	12.0	13.0	13.5	13.5	4560	
Estimates of σ (= $^1/_b$)							
1.5	0.5	-1.0	0	-0.5	0	900	R.W.
1.5	2.5	2.0	1.0	1.5	2.0	4560	
2.0	2.0	4.0	1.0	1.0	2.0	900	L.M.
0	1.5	0.5	1.5	1.0	-1.5	4560	
2.0	2.0	2.0	4.0	2.0	0	900	K.R.
1.0	2.0	2.0	4.0	3.0	1.0	4560	

All values quoted in units of 5.8 microseconds

TABLE 6.10 : ANALYSIS OF VARIANCE - WETHERILL ESTIMATES

Effect	m estimates: Mean Square	m estimates: Sign.Level	s estimates: Mean Square	s estimates: Sign.Level
Time-Order (A)	0.72	-	1.32	-
Bandwidth (B)	90.3	.995	0.11	-
Subjects (C)	670.8	.995	4.33	.95
Interactions:- A x B	0.92	-	0.74	-
B x C	146.7	.995	7.53	.99
A x C	3.39	.95	1.40	-
Residual	0.754	-	0.811	

FIG. 6.6 : WETHERILL STRATEGY-ESTIMATES OF 30% AND 70% POINTS



As before, variations in m were very large between occasions and since tests on different subjects and with different bandwidths were carried out on separate occasions, these factors appeared to be significant in m.

The Wetherill strategy was of great value in that efficient and relatively unbiased estimates of the difference limen were obtained. The results showed no significant effect of time-order on the estimated lateralization ability ($= {}^{1}/s$), even when a marked trend in m was observed - see Fig.6.6. More convincing evidence of the stability of s was obtained in the subsequent experiment (C1.).

Judging by the comments of listeners after the experiments, the subjects were not consciously aware of any sequential strategy being used. Subjective assessments, however, are notoriously unreliable and an experimental check on a listener's performance under different strategies was desired. Since the same effects were observed in this experiment as in experiment B4, it seemed unlikely that the results were significantly dependent on the experimental technique.

6.8 Differences in Performance for the Various Strategies

It was of great interest to determine whether the experimental strategy had any noticeable effect on a listener's performance. The comparison was made over four subjects using, for the most part, data from previous experiments. Where necessary, however, additional tests were carried out. Since one subject had left the country at the time, the design was incomplete in one cell. Yates's method of compensating for a missing plot was used (Davies 1956 p.159).

The factors tested were:

Strategy (3 levels; quasi-random, Up-and Down, Wetherill 30 and 70% points)

Subjects (4 levels; G.S., R.W., K.R. and L.M.).

Low-pass noise of 4560 c.p.s. bandwidth was used throughout. The results are quoted in Tables 6.11 and 6.12.

The comparison was carried out using a common estimation procedure (maximum likelihood) for all the data since the object of the analysis was to detect changes in the listener's performance and not statistical differences between the estimates. Since the probit technique implicitly assumes that μ is constant, observed differences in b could be due either to a change in lateralization ability or to the effect of a gradual drift in medial plane position. A more effective comparison was not possible since one of the strategies (the quasi-random) does not cater for variations in medial plane position.

Although the observed variation in b between the different strategies did not quite reach the 90% level,

TABLE 6.11 : EFFECT OF SEQUENTIAL STRATEGY

- b ESTIMATES.

Subjects				Strategy
G.S.	R.W.	L.M.	K.R.	
.169	.335	.531	.054	Up-and-Down
-	.272	.127	.591	Wetherill
.261	.811	.678	.659	Quasi-Random

1 unit = 5.8^{-1} microseconds^{-1}.

Note: Probit estimates were used in order to provide a common comparison; the method of placing the observation followed the given strategy.

TABLE 6.12 : ANALYSIS OF VARIANCE

- SEQUENTIAL STRATEGY, b ESTIMATES.

Effect	Mean Square	Significance Level
Strategy	.153	nearly .9
Subjects	.067	-
Residual	.0461	-

there was a significant difference in b between the quasi-random case and the two sequential strategies. A comparison between the m estimates was not carried out since it would have been of little value. The tests were performed on separate occasions - a factor known to be highly significant in m (but not b). A notable difference between the various techniques, however, was that drifts in m were more readily observed using a sequential strategy. Whether this was due to the greater efficiency of the sequential strategy or whether the listener's performance was more stable with the quasi-random strategy, is an open question.

Subjectively, a characteristic feature of the quasi-random selection technique was that every now and again an "easy" judgement was required of the listener. With either of the sequential strategies the listener's task was consistently difficult. It was therefore possible that the listener could have adapted to the experimental situation, hence performing differently depending on the strategy. Since a slight change in performance was in fact observed, this raised the important and difficult question: which strategy should be used?

The quasi-random technique has the advantage that in certain respects it is more general and that it can provide a fairly reliable check on the shape of the lateralization curve. This apparent generality, however, is deceptive in that the range of δT must be decided on beforehand. A decision which is either arbitrary, or based on the results of previous experiments or other prior knowledge, which in essence is the beginning of a sequential strategy.

Among the many advantages of a sequential strategy are greater efficiency, comparative lack of bias and greater versatility. If a listener's performance does depend on the average difficulty of a test the sequential strategy can at least ensure that the test be of constant average difficulty. This is perhaps less arbitrary than selecting the range of δT beforehand with little knowledge of the listener's ability. The philosophy of matching test difficulty to the listener's performance as the test progresses is, in the author's opinion, a good one and may well lead to more reliable test procedures.

The great danger associated with sequential techniques is the possibility of the listener recognizing a pattern in the presentations. The use of more than one strategy with associated presentations interleaved at random, however, will effectively conceal any obvious regularities in the test procedure without sacrificing the statistical advantages of the technique. A suitably randomized sequential strategy was therefore recommended for subsequent experiments.

6.9 Summary and Conclusions

It was shown that two independent but statistically identical random noise sources were indistinguishable from a single source of twice the power, provided the two sources were closely spaced; i.e. $\delta T_1 - \delta T_2 < 5$ units. If the sources were widely spaced separate noise images were perceived by the listener although the estimated degree of separability was small. This effect was demonstrated using three noise sources so that the separate images could be correctly identified.

The results were interpreted in terms of a simple crosscorrelation model. Quantitative aspects of the model were tested by measuring the effect of noise bandwidth on lateralization ability and ability to detect one from two noise sources. The results were in reasonable agreement with the model. An unexpected observation was that bandwidth appeared to have a negligible effect on lateralization ability - at least for low-pass noise with a cut-off frequency in the range 900 to 4560 c.p.s.

A distinctive feature of the proposed model is that the variability of a listener's responses is ascribed to two quite separate sources - a variance component typical of the listener and a component related to the shape and variability of the running correlation function. In this way differences between subjects can be accounted for as well as general trends.

Sequential strategies were employed and the results were encouraging. Unlike the ambiguous result of Series A a highly significant trend in m on certain occasions was observed. This finding was presumably due to the considerable efficiency of the sequential estimation technique, although the strategy itself may have influenced the listener's performance. No significant variation in performance was observed between the two sequential strategies, but a slight difference for the quasi-random selection technique was noticed. Of the various estimation techniques a suitably randomised sequential strategy was considered to be most suitable. Apart from the greater efficiency of estimation the strategy would enable gradual drifts in medial plane position to be tracked; furthermore, the average difficulty of a test could be matched to the subject's performance. With the quasi-random technique the range of the physical variate must be selected beforehand, a decision which in many cases would be based on insufficient information.

CHAPTER 7 USE OF SEQUENTIAL STRATEGY IN MEASURING SEPARABILITY

7.1 Introduction

In the preliminary experiments (Series A) it was pointed out that any shift in the position of the medial plane during the course of a test would have an adverse effect on the proposed measure of separability since,

a) the estimated slope, b, of the lateralization curve would invariably be decreased and it would not be possible to randomise out this effect as with other nuisance parameters, and

b) the reliability with which a sound image could be placed in the medial plane was open to question.

The results of the preceding experiments (Series B) suggested that the difficulties associated with an unstable medial plane might be considerably reduced by adopting a sequential strategy. In this section an attempt was made to measure the degree of separability of one of two physically similar speech sounds using the highly efficient Wetherill strategy. By so doing it was hoped to demonstrate a convenient and practical method for measuring the degree of separability while at the same time considering an extremely interesting application of the proposed measure; namely the relative separability of two speech sounds which are physically almost identical, differing only in semantic content.

7.2 Experiment C1. Separability of Two Physically Similar Voiced Messages

7.2.1 Purpose

To measure, using the Wetherill sequential strategy, the degree of separability of typical voiced speech in the presence of an interfering speech sound which is virtually identical in all respects other than semantic content, and to investigate the effect on this measure of a relative intensity difference between the two sounds. A secondary aim of the experiment was to compare two versions of the Wetherill strategy in carrying out the above measurements.

It should be noted that Egan, Carterette and Thwing (1954) have carried out a similar investigation in which the relative articulation of the wanted speech message was measured. A sharp change in the slope of the articulation function was observed when the intensities of the two speech sounds were equal (Fig.7.1); it appeared as if in this region a slight lowering of the intensity of the wanted sound led to an improvement in articulation. This is a particularly interesting result since it represents one of the few cases where a decrease in a physical variate such as would normally impede perception, is in fact of assistance.

Usually it has been found that those physical factors which aid lower order processes, (e.g. the detection of stimuli as sounds) also reflect an improvement at the higher levels, (e.g. the perception of sounds as messages). This is presumably why Fletcher was so successful in predicting the intelligibility of speech in noise from simple masking data (Fletcher 1953). The formulation of a separate sound image for one of several simultaneous sound sources is an essential stage in the separation process and it was of interest to check if a similar connection existed between the degree of separability and relative intelligibility (or articulation).

7.2.2 Description

Several paired recordings of a male voice reading either a novel or a scientific text were used. The recordings were made under identical conditions using the same speaker. The speech was at a normal rate but with as few pauses as possible. The average level, including pauses, of each of the six ½-hour recordings was measured on a true r.m.s. meter (Solartron JM1067) with an averaging time of well over a minute;

long term fluctuations in average level were found to be within $\pm$ 1 dB. A pure tone of equivalent mean power preceded each recorded passage and level settings during the course of the experiment were carried out using this test tone. Both speech sounds were presented simultaneously and the listener was required to lateralize the "scientific" speech about the medial plane. The apparent position of the unwanted speech (the novel reading) was varied at random but always in the vicinity of the medial plane. The level of the wanted message was 55 dB above threshold for all tests; the level of the unwanted speech was fixed for each test but varied between tests over a range from 51 dB to 67 dB above threshold in steps of 4 dB. A control set of tests using only the scientific readings was also carried out.

Wetherill's sequential strategy, as described in section 4.3.2, was employed. Two points on the curve were estimated, the allocation of presentations for estimating each point following a random sequence. The percentage points estimated were either the 30% and 70% points or the 16% and 84% points, part of the experiment being a comparison between the different strategies. For the 30 and 70% points a change in the direction of increasing δT was not made until two consecutive positive responses had been obtained; for the

16 and 84% points four consecutive, identical responses were required. Increments of 2 units in δT were used. Each test started with a trial run using the simple Up-and-Down method; the Wetherill strategy was adopted, without the listener's knowledge, after a reasonable estimate of the 50% point had been obtained. A test was terminated after sufficient runs for estimating both percentage points had been obtained; at least 7 runs for each of the 16 and 84% points and at least 9 runs for each of the 30 and 70% points. The initial Up-and-Down estimates usually involved 4 runs. On average it was found that a test required 60 to 80 presentations and lasted about 20 minutes.

A standard factorial design was used, the following factors being investigated:

Level of Interfering Speech (6 levels; absent, -4, 0, +4, +8, +12 dB relative to wanted message)

Subjects (5 levels; L.M., T.M., K.P., K.R. and R.W.)

Time-Order (3 or 4 levels; depending on strategy employed).

The experiment was repeated for each of the two versions of the Wetherill strategy. Since the data was obtained sequentially a check on time-order was possible, but the

FIG. 7.1 : RELATIVE ARTICULATION FOR ONE OF TWO PHYSICALLY SIMILAR MESSAGES. - From Egan et al (1954)

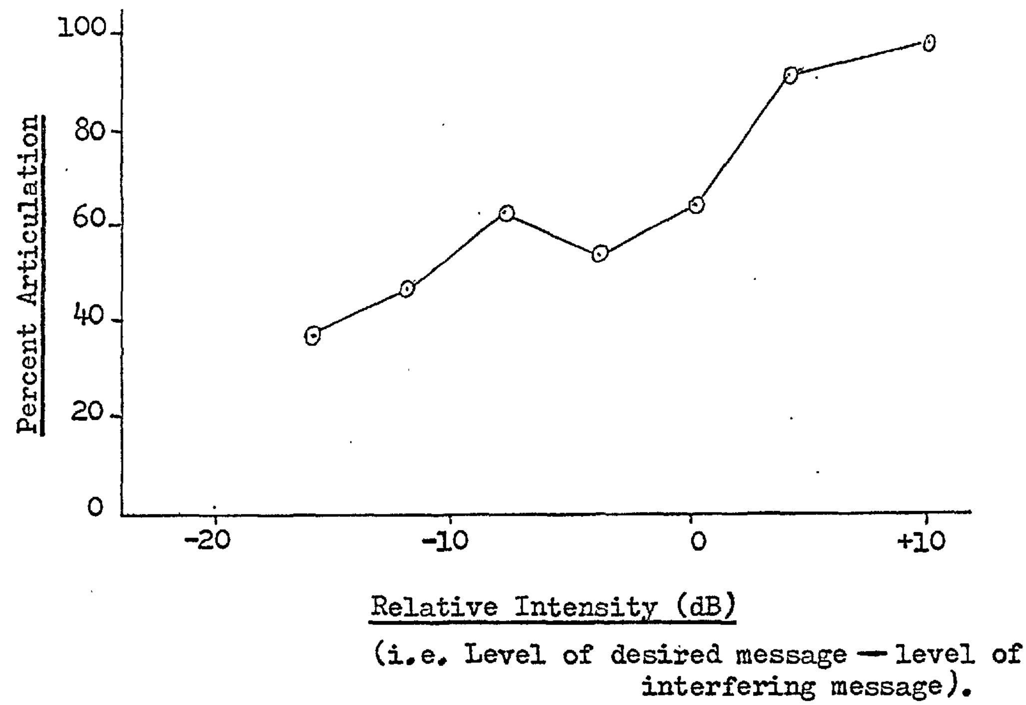
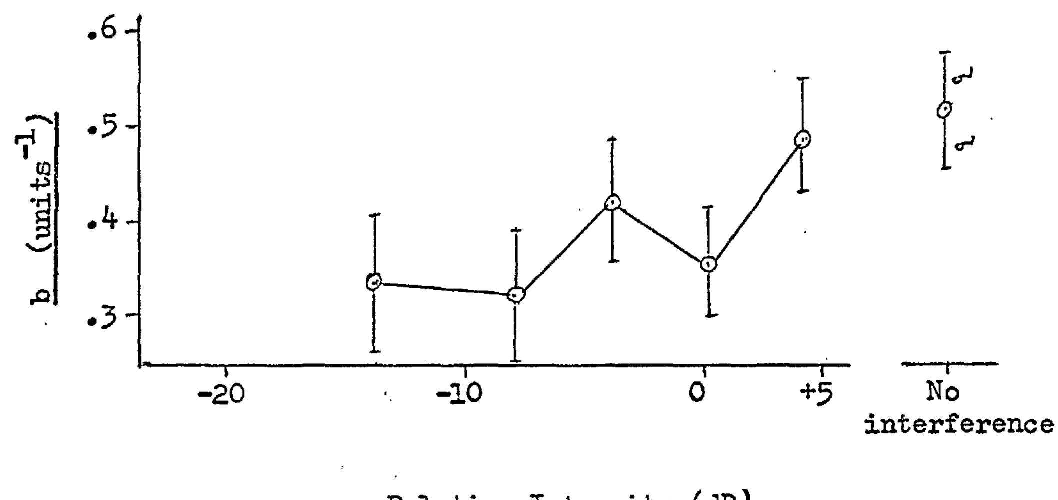


FIG. 7.2 : LATERALIZATION ABILITY FOR ONE OF TWO PHYSICALLY SIMILAR MESSAGES.



Note: Intensity of desired message was held constant for all tests; in Egan's study intensity of interference was held constant.

TABLE 7.1 : VARIATION IN s, EXPT. C1.

Subject	Relative Level of Interference	Time-Order: 30 & 70% Strategy 1	2	3	4	16 & 84% Strategy 1	2	3
L.M.	Absent	0	1	3	1	3	3	2
	-4dB	-1	2	2	2	1.5	1	1
	Odb	2	2	2	3	3	3	1.5
	+4dB	4	2	2	1	1.5	1.5	2.5
	+8dB	4	4	4	4	2.5	2	1
	+12dB	3	1	2	5	4	3.5	5
T.M.	Absent	1	1	2	5	3.5	2.5	0.5
	-4dB	5	3	3	2	2	1.5	2
	0 "	6	7	5	3	2.5	3	2.5
	4 "	2	2	4	4	3.5	2.5	4.5
	8 "	5	9	8	4	2	3	3
	12 "	1	4	2	2	4	2.5	1.5
K.P.	Absent	2	3	4	5	2	2.5	2
	-4dB	3	2	4	9	1.5	3.5	3
	0 "	5	3	4	3	2	3	4.5
	+4 "	4	6	6	7	3	2	3
	+8 "	2	4	4	1	6	4.5	1
	+12"	5	2	6	4	3.5	3	3
K.R.	Absent	-1	0	2	1	2	1	1.5
	-4dB	3	0	4	4	2.5	2	1
	0 "	2	0	0	0	5	5.5	5
	+4 "	2	1	2	-1	2.5	1	1.5
	+8 "	0	0	4	4	1.5	2.5	3.5
	+12"	3	2	2	3	2.5	2.5	1
R.W.	Absent	3	1	3	2	1	1	1
	-4dB	1	2	1	1	0	0.5	1
	0 "	2	2	2	2	1	1	1.5
	+4 "	2	1	-1	1	0	1	1.5
	+8 "	0	2	2	2	3	1	1.5
	+12"	0	2	1	-1	2	2	3

1 unit = 5.8 microseconds.

TABLE 7.2 : ANALYSIS OF VARIANCE IN s

- Experiment C1

Effect		30 & 70% Strategy Mean Square	Sign. Level	16 & 84% Strategy Mean Square	Sign. Level
Time-Order	(A)	2.88	-	0.50	-
Level of Interference	(B)	4.13	nearly 0.9	4.39	.995
Subjects	(C)	37.43	.995	7.03	.995
Interactions:-	A x B	2.25	-	0.75	-
	B x C	5.57	.995	1.96	.95
	A x C	1.70	-	0.29	-
Residual		2.13		0.90	

TABLE 7.3 : VARIATION IN m, EXPT. C1.

Subject	Relative Level of Interference	Time-Order: 30 & 70% Strategy 1	2	3	4	16 & 84% Strategy 1	2	3
L.M.	Absent	2	1.5	0.5	0.5	-1	-1	-2
	-4dB	-1.5	-2	0	0	7.5	7	7
	0 "	2	3	4	4.5	0	1	-0.5
	+4 "	-1	0	0	-0.5	3.4	4.5	6.5
	+8 "	-2	-4	-6	-6	0.5	1	1
	+12"	4.5	5.5	6	4.5	1	1.5	2
T.M.	Absent	6.5	5.5	6	5.5	-8.5	-10.5	-8.5
	-4dB	-0.5	-0.5	0.5	1	8	7.5	7
	0 "	0	1.5	2.5	3.5	-1.5	-1	-0.5
	+4 "	2	1	2	2	1.5	0.5	-0.5
	+8 "	0.5	0.5	3	7	-8	-10	-12
	+12"	-11.5	-11	-10	-8	-4	-5.5	-5.5
K.P.	Absent	-8	-10.5	-11	-11.5	-6	-8.5	-8
	-4dB	-6.5	-6	-7	-9.5	10.5	10.5	12
	0 "	-5.5	-6.5	-7	-6.5	10	13	13.5
	+4 "	4	5	7	8.5	11	12	13
	+8 "	-12	-13	-13	-14.5	-8	-8.5	-11
	+12"	-2.5	-3	-2	-3	8.5	9	10
K.R.	Absent	6.5	6	9	8.5	-8	-11	-11.5
	-4dB	3.5	3	4	4	7.5	8	9
	0 "	-9	-10	-8	-8	-5	-5.5	-8
	+4 "	11	11.5	11	14.5	-7.5	-9	-9.5
	+8 "	-6	-6	-6	-7	11.5	13.5	14.5
	+12"	5.5	5	4	4.5	11.5	13.5	16
R.W.	Absent	-4.5	-5.5	-5.5	-5	5	5	5
	-4dB	-2.5	-3	-3.5	-4.5	4	4.5	6
	0 "	5	5	5	6	-5	-5	-6.5
	+4 "	1	1.5	1.5	1.5	-4	-4	-5.5
	+8 "	4	5	5	5	+1	-1	-0.5
	+12"	4	5	5.5	5.5	5	5	5.5

1 unit = 5.8 microseconds

TABLE 7.4 : ANALYSIS OF VARIANCE IN m

– Experiment C1

Effect		30 & 70% Strategy Mean Square	Sign. Level	16 & 84% Strategy Mean Square	Sign. Level
Time-Order	(A)	2.2	-	0.18	-
Level of Interference	(B)	125.3	.995	291.7	.995
Subjects	(C)	226.9	.995	132.6	.995
Interactions:	A x B	1.1	-	1.7	-
	B x C	141.5	.995	156.0	.995
	A x C	2.0	-	0.92	-
Residual		1.32		1.41	

number of estimates per test were limited to 4 for the 30 and 70% estimates and 3 for the 16 and 84% estimates. Each test was carried out on a separate occasion. Headphones, channels, recording tapes, ordering of tests and other possible sources of error were varied at random between tests.

In this investigation, unlike any of the others, the presentations were time-limited. The reason was that with no time-limiting the listener could, if desired, wait for a pause in the interfering speech before making a decision, in which case the effect of the interference would be almost negligible. A presentation duration of 5 seconds was selected since this allowed sufficient time for the listener to identify the sounds; pauses within this period were relatively infrequent.

7.2.3 Results and Discussion

The results with the associated analysis of variance are quoted in Tables 7.1 to 7.4. The estimate s rather than b ($= 1/s$) was used since it was obtained directly using the Wetherill technique.

Lateralization ability was noticeably impaired by the presence of the interfering sound. The effect was most marked when the two sound sources were either equal in intensity or when the interfering speech was very much louder. The results are summarised in Fig.7.2.

Although an effect similar to that reported by Egan et al (1954) was observed, the results were disappointing in that the observed loss in lateralization ability was comparatively small compared to the experimental error. It may be noted that when both sound sources were equal in level the observed loss in lateralization ability was presumably due in part to a loss in discriminability, i.e. the listeners claimed that with no loudness disparity it was sometimes difficult to identify the voices. Egan et al attribute the relative loss in articulation for the equal loudness situation to the same cause.

Judging by listeners' impressions, as recorded after the tests, it appeared that only the most prominent features of the desired speech sound were used by the subjects for the lateralization task. In view of the considerable short-term fluctuations in the intensity of running speech, it would require a particularly loud interfering speech sound to swamp all prominences in the wanted message. This property of speech has been encountered in a variety of forms, (e.g. the auditory threshold for speech has at least three distinct levels; as a sound, as a speech-like sound and as intelligible speech) and it would appear that lateralization ability, and hence degree of separability, are measures which depend largely on prominences in the speech waveform and not on its average level.

As an empirical method of predicting speech intelligibility, the measure of separability was considered to be impractical owing to the relatively small size of the observed effect. It might be remarked, however, that the method at least indicated the importance of a loudness disparity in separating and discriminating between very similar speech sounds. Fletcher's method of predicting intelligibility from threshold masking could not possibly anticipate this effect.

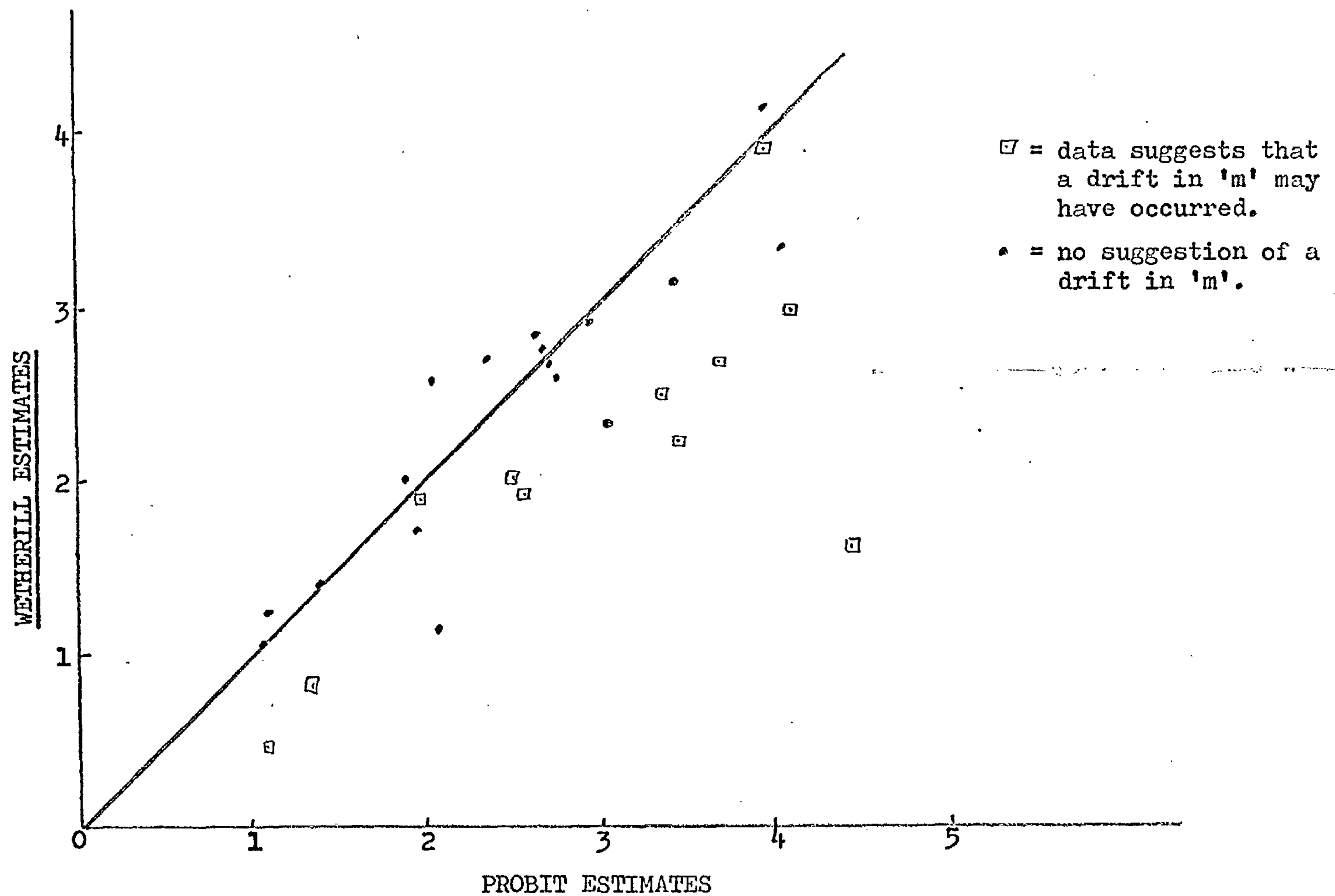
In comparing these results with the data of Egan et al it should be noted that in the articulation tests, average level of the wanted message was varied. In this study the level of the interfering sound was the controlled variable; the level of the wanted message was held constant in keeping with the definition of degree of separability (i.e. the loss in lateralization ability for the desired sound given some form of interference). For nearly equivalent speech levels, which was the region of greatest interest, Egan's data were of direct relevance.

As in all the previous experiments, differences between subjects were found to be large. In addition, a significant interaction between subjects and the level of the interfering sound was observed. This suggested that certain subjects (notably R.W.) were less sensitive to the interfering sound than others.

A matter of great interest was the effect of time-order. It was apparent even while performing the experiment that on certain occasions there was a marked drift in the estimated position of the medial plane. Variations in m are quoted in Table 7.3 and the resulting analysis of variance in Table 7.4. The results which indicated a drift were similar in form to that shown in Fig.6.6 of the previous chapter. In all cases it was found that despite the marked variation in m the distance between the two estimated percentage points remained virtually constant. Since consecutive estimates were directly available from the data it was possible to test for any regular trends in s during a test. No significant effects were observed, whether or not a large drift in m had occurred. This was a particularly encouraging result in that it demonstrated the practical value of the Wetherill strategy in compensating for drifts in medial plane position.

A comparison between the Wetherill estimates of the difference limen and those obtained by probit analysis (using the same data) is shown in Fig.7.3. The two estimates were usually very similar and hence tended

FIG. 7.3 : COMPARISON BETWEEN WETHERILL AND PROBIT ESTIMATES OF σ



to lie on a line of 45° slope through the origin. In certain cases, however, the probit estimates of σ were slightly greater than the corresponding Wetherill estimates. In all these cases a drift in m was observed. Fig.7.3 provides a good idea of the errors which can arise if probit analysis is applied to data where the position of the medial plane has drifted.

Two versions of the Wetherill strategy had been tried out in order to

a) check whether the listener's performance was affected by variations in the strategy, and

b) to determine which version was more convenient to use in practice.

The results showed no significant variation in lateralization ability or stability of the medial plane between the two versions of the strategy. The experimental error for s was found to be less with the 16 and 84% strategy than with the other; on the other hand the 30 and 70% estimates required fewer presentations on average. A comparison between the efficiency of the two strategies has been drawn up in Table 7.5. Since each test was limited in duration, the comparison was based on the error variance for a test of fixed size. It was found that estimating the 30 and 70% points led to a more efficient estimate of the midpoint whereas

TABLE 7.5 : COMPARISON BETWEEN ERROR VARIANCES

EXPT. C1

Estimated Parameter	Strategy	Observed Variance	Av.no.of presentations per test	No. of estimates per test	Predicted error variance per 10 prstns.
μ	30 & 70%	1.32	60	4	2.0
	16 & 84%	1.41	76	3	3.7
σ	30 & 70%	2.13	60	4	3.2
	16 & 84%	0.90	76	3	2.3

Variance values quoted in units of (5.8 microsecs.)^2.

estimating the 16 and 84% points led to a correspondingly more efficient estimate of the difference limen. The 30 and 70% strategy was found to be easier to use in practice and was also more efficient in compensating for drifts in medial plane position.

7.3 Implications of Experiment C1.

The previous experiment demonstrated that although a loudness disparity between two physically similar speech sounds influenced the degree of separability, the effect was not very great. This was possibly because the fluctuating nature of the speech sound allowed for reliable lateralization on the basis of only a few prominent peaks in the wanted message. If so, the separability of those sounds which exhibit large, short-term fluctuations (such as speech) should be fairly invariant for a wide range of interfering sounds of a similar nature. It would also be expected that the degree of separability for a wide variety of competing speech messages would be much the same.

The estimation strategy used in experiment C1 was notably superior to any technique previously employed and it would appear to be the answer to the difficulties encountered in the preliminary experiments. The estimates of lateralization ability were found to be reasonably independent of any trends in m; furthermore, the nature of the strategy was such as to allow the midpoint to be reliably estimated during the course of a test. Hence it would be possible to place a sound image in the medial plane with reasonable accuracy despite a gradual drift in its position. To illustrate how this approach could be employed, a pilot experiment (C2) was carried out using two pairs of sound sources which were indiscriminable, but which could possibly yield separate images.

7.4 Experiment C2. Placing a Sound Image in the Medial Plane using a Sequential Strategy

7.4.1 Purpose

To demonstrate how a sound image may be placed in the medial plane using a sequential strategy.

7.4.2 Description

Two pairs of sound sources were used; a pair of statistically identical random noises and two similar recordings of backward speech. The latter were tape recordings played backwards, and were made by the same speaker under the same conditions, reading a similar text. Each pair of sound sources was matched in level, namely 55 dB above threshold. Both the backward speech and the two noise sources were indiscriminable, but it was expected that the speech would give rise to separate though indiscriminable sound images.

Presentations using one or two sound sources were interleaved during the course of a test. The listener was instructed to lateralize the outer sound image about the medial plane if two images were perceived. If a single image was heard then that image was to be lateralized. Estimates of the position of the medial plane were based on the most recent judgements for a single source. The sequential strategy employed was that for estimating the midpoint directly, i.e. the Up-and-Down method. A trial run estimating 30 and 70% points was also carried out. In this case mid-point estimation was less efficient and required about 8 presentations on average before an estimate was obtained. One subject was used. The factors investigated were:

Type of Sound Source	(2 levels; Noise, Backward speech)
Number of Sources	(2 levels; 1 or 2 sources).

7.4.3 Results and Discussion

The results are quoted in Table 7.6. The data for the noise sources were as expected, namely a marked decrease in estimated lateralization ability for two noise sources was observed. There was, however, no significant difference in b whether one or two backward speech sources were presented - i.e. the associated sound images appeared to be totally separable. The technique seemed promising and may well be of use in subsequent investigations on separability.

TABLE 7.6. Estimated Lateralization Ability for White Noise and Backward Speech - Subject G.S.

Source	b	Standard error of estimate
Single Noise Source	.216	.069
Two Noise Sources	.121	.052
Backward Speech; 1 Voice	.342	.111
Backward Speech; 2 Voices	.475	.152

1 unit = 5.8^{-1} microseconds^{-1}.

7.5 Summary and Conclusions

A highly efficient sequential strategy was employed in order to measure the effect of a loudness disparity on the degree of separability of one of two physically similar speech sounds. The observed effect was not very large, but was the same in form as that reported by Egan et al (1954) for relative articulation under similar conditions. The accuracy with which a listener could lateralize one speech sound in the presence of another more intense sound may well have been due to the listener using only prominent peaks of the wanted speech for the lateralization task. This suggested that lateralization ability (and hence relative separability) for sounds which fluctuate widely in level would be relatively independent of similar interfering sounds.

The estimates of the difference limen were found to be reasonably stable despite marked drifts in m. This was attributed to the efficiency of the sequential estimation strategy in tracking and compensating for gradual variations in medial plane position. The use of a sequential strategy in placing a sound image in the medial plane was also demonstrated. This technique, or variations of it, was recommended for use in subsequent investigations.

CHAPTER 8 SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH

8.1 Summary

The property of separability and its method of measurement have been defined. The technique was based on measuring the change in a listener's ability to lateralize a sound about the medial plane when a second, interfering sound was introduced. If no noticeable change occurred, then the first sound was said to be totally separable from the second; if a change was observed then the loss in lateralization ability was used as a measure of the degree of separability. Headphone listening with interaural time delay as the physical correlate of lateral position was used. In order that the listener should be able to identify the sound images, (should they be separable), it was proposed that one image be placed in a reference position, namely in the medial plane.

A preliminary series of experiments showed that lateralization ability (and hence changes in lateralization ability) could be measured with considerable precision. The results also indicated that the estimated position of the medial plane (in terms of δT) may possibly have drifted during some of the tests. This finding was particularly disconcerting for two reasons;

a) instability of the medial plane could lead to a marked decrease in the estimate of lateralization ability,

b) doubts were raised as to whether a sound image could be reliably placed in the medial plane.

Severing all connection with the medial plane would be one way of avoiding these difficulties, but this could only be done by introducing an arbitrary reference image; which in itself raised serious objections. Furthermore, a pilot study on lateralization ability for one sound image about another similar image (male and female whispered speech were used), showed that there was a marked decrement in image resolution with distance from the medial plane. Placing both sound images away from the medial plane would therefore introduce an additional nuisance parameter; average distance from medial plane. Other strategies, such as compensating for shifts in medial plane position during the course of a test, seemed more appropriate and hence were pursued.

The preliminary experiments also provided useful information on the statistical properties of the data. The assumption of a cumulative normal distribution for the lateralization curve was found to be justified. An efficient allocation of data for complex experiments was also determined.

The second series of experiments (Series B) considered several of the points raised by the preliminary series. The ability of a listener to discriminate between one and two similarly placed, statistically identical noise sources was measured and shown to be dependent on the bandwidth of the noise. On the other hand, lateralization ability for a single noise was found to be reasonably independent of bandwidth - at least for low-pass noise with a cut-off frequency in the range 900 to 4560 c.p.s. For three widely spaced noise sources separate noise images were perceived, although the degree of separability was small. The results were interpreted in terms of a simple cross-correlation model and reasonable quantitative agreement was found between observed and predicted values.

Sequential estimation techniques were employed in the above experiments and the results were encouraging. Unlike the ambiguous results of Series A, a highly significant trend in m on certain occasions was observed. This finding was presumably due to the considerable efficiency of the sequential strategy although the use of a sequential technique may itself have influenced the listener's performance. It was argued that of the various possible estimation techniques, a suitably randomised sequential strategy was to be preferred.

Apart from the greater efficiency, it also enabled gradual drifts in medial plane position to be tracked, thus providing an answer to the difficulties encountered in the preliminary experiments.

The newly developed Wetherill strategy was used in an investigation of the relative separability of physically similar, voiced messages. The results showed an effect similar to that observed by Egan et al (1954) on relative articulation under much the same conditions. The effect, however, was not very marked and hence not of great practical significance. The estimation technique was found to be both efficient and versatile and gave consistent results despite occasionally large drifts in m.

In conclusion it may be said that the proposed measure of separability appears to be practically feasible provided a suitably randomised sequential strategy is used. The Wetherill strategy is recommended for this purpose.

8.2 Suggestions for Further Research

Several interesting points were raised during the course of the investigation. These and other ramifications are described below.

8.2.1 On the Measure of Separability

The method of measuring separability has been shown to be practically feasible. Using the technique two limiting cases have been observed, namely two voiced or whispered speech sounds which were both separable and discriminable and two statistically identical noise sources with similar interaural delays which were totally inseparable. Whispered speech may be simulated approximately by operating on random, white noise in various ways (e.g. spectrum shaping, selective modulation) and it would be of great interest to measure relative separability at various stages in this process.

Situations involving more than two sound sources would appear to be worthy of further investigation. In experiment B2, for example, a marked decrement in lateralization ability was observed for one of three separate noise sources. Could it be that lateralization ability is adversely affected by more than one interfering sound image? A multiplicity of spatially separate sound images has been shown to have an adverse effect on other, more complex, listening tasks. Treisman (1961), for example, has shown that shadowing one of three spatially separate sound images was less reliable than if two spatially separate images occurred - albeit three voices were still heard. This effect is of direct relevance to the important practical situation where a listener is faced with a barrage of voices emanating from different directions.

The use of three sources would also be of value in allowing the listener to identify the sound images without having to fix one image in a reference position, with its attendant difficulties. Using this technique various aspects of the proposed model could be examined. In particular it would be of interest to determine the minimum distance between noise sources (along the δT axis) for distinctly separate sound images to be heard.

8.2.2 Extensions to the Correlation Model

The Cherry-Sayers model was concerned primarily with lateralization; in this study a development of the model was used to predict the separability of multiple noise images. Further extensions are possible. For example, the fusion process itself may be related to the presence of a peak in the correlation pattern, and loudness or threshold masking may correspond to the height of a correlation peak above the surrounding noise.

Two assumptions were implicit in the proposed correlation model; linearity and time-invariance. Owing to the large number of variable parameters in the model these assumptions were difficult to check. On the other hand a comprehensive model has potentially more degrees of freedom and an extensive study of the correlation model as such could provide the necessary data for testing these assumptions. Physical variables which may be easily controlled by the experimenter are;

a) shape of the autocorrelation function (e.g. by means of linear networks),

b) degree of correlation between inputs (e.g. by adding uncorrelated noise to each input),

c) differences between interaural time delays,

d) use of random or deterministic signals.

Central to any correlation model is the assumption of a time window or weighting function. An important aim of the proposed investigation would be to determine the shape and effective width of such a time window, and also whether there is any evidence of a change in the size or shape of the window for different types of stimuli.

The limits of binaural fusion need to be fully explored This could well be done in the context of a correlation model. For example, a recent study (Blodgett, Wilbanks and Jeffress 1956) has shown that for low-pass noise the critical interaural delay for the breakdown of fusion decreases with noise bandwidth; this could be interpreted in terms of the greater spread and variability of the correlation peak for noise of small bandwidth.

8.2.3 Binaural and Interaural Effects

It has been shown by Treisman (1963) that masking can occur both peripherally and centrally. The relative role of each type of masking may be worth investigating using the following artifice. A stimulus is switched rapidly from ear to ear such that a single sound image is heard (Cherry 1953). A second stimulus is introduced which is also alternated between the ears at the same rate. The relative phase of the switching function may be adjusted such that the waveforms of each stimulus either overlap at each ear or are interleaved. When the signals overlap both peripheral and central masking should occur, when the waveforms are interleaved mainly central masking should occur. In this way a measure of the relative importance of peripheral masking may be obtained.

An interesting experimental finding was the observation (section 6.3.4) that several noise sources led to multiple sound images, even if interaural intensity differences rather than time differences were used. It is not immediately obvious how this effect could be explained in terms of the current theory that intensity differences are transformed into time differences as a result of more rapid neural firings for stimuli of greater intensity. Further experiments on competing stimuli using interaural intensity differences may throw more light on this trading process - if it exists.

The useful exploitation of binaural phenomena is a matter of direct interest to engineers, and the following method of improving channel efficiency may be worth pursuing.* Consider several sound sources and a corresponding number of listeners, but only two transmission channels. The signal from each source is alternated between the two channels at a rapid rate (switching period $<$ 0.1 seconds); in addition, a large interchannel time delay ($>$ 20 millisecs.) characteristic of each source is introduced. The delays are sufficiently large so as to ensure that if the signals in the two channels were to be applied dichotically no fused image would be heard. On the other hand if a complementary inter-channel delay characteristic of one of the sources is introduced, a central, fused image corresponding to that sound source will be heard, with the remaining sounds appearing as a babble of interrupted, speech-like sounds at each ear. Since intelligibility is improved by spatially separating wanted and unwanted messages this scheme should improve

---

* Pilot equipment along the lines suggested is now under construction in the Department of Electrical Engineering, Imperial College of Science and Technology, London, S.W.7.

listening efficiency in the multiple source situation. A slightly different scheme involving several channels has been described by Pollack (1959), and one for spatially separating message from noise in a dual single-sideband communication system by Reynolds, Voelcker and White (1961).

8.2.4 Application of the Sequential Estimation Technique

The Wetherill strategy has many advantages of both a statistical and psychophysical nature. It is particularly suited for estimating difference limens and it is hoped that it will find wide application in psychophysics generally. The technique may be of some value in the current controversy on the shape of the psychometric function.

A possible and perhaps very useful application of the technique would be in the diagnosis of loudness recruitment. Direct measurement of the loudness function for each ear is both tedious and time-consuming and several alternative methods of diagnosis have been suggested (e.g. see Egan 1954). For example, one method would be to use a Bekesy audiometer for measuring the auditory threshold; if a small fluctuation in the trace is observed then this is evidence of an extremely steep loudness function and hence an indication of recruitment. This technique is not very reliable in that it involves the reaction time of the subject in controlling the device. The Wetherill strategy would provide much the same information but with greater precision.

The philosophy of matching the average difficulty of a test to a subject's ability is, in the author's opinion, a good one and may well lead to an improvement in the repeatability of subjective tests. It may be worthwhile investigating possible extensions of the sequential strategy to the general problem of scaling.

8.2.5 Computer Techniques

The use of a computer for carrying out analysis-of-variance computations is of great value, but suffers from one serious drawback; the preparation and testing of a programme is time-consuming. Unfortunately the range of possible analyses of variance is very wide and at present there is no general programme for covering all possible cases, although work in this direction is being pursued (Yates, Gower and Simpson, 1963). The suggestion is made that the simple programme described in Appendix A2 be used as a basic routine and the computational information for any given analysis of variance be fed, in coded form, with the data. This would require a library of special "headings" for the different types of data, but at least the headings would be relatively easy to prepare and the library could be extended at will.

8.3 Conclusions

The primary purpose of this study has been to investigate the practical aspects of a proposed measure of sound-image separability. The conclusion reached was that the technique was practically feasible, but with certain reservations; a highly efficient estimation technique must be employed in order to cope with gradual drifts in medial plane position. The recently developed sequential strategy of Dr.B.Wetherill was recommended.

In the course of the experiments one or two interesting digressions were made. In particular, several experiments with separate but statistically identical noise sources were carried out; the results were interpreted in terms of a simple crosscorrelation model. The experiments, however, provoked more questions than they attempted to answer. The author hopes that some of the points raised may be of value to others working in this and related fields.

APPENDIX A1 THE CORRELATION MODEL AND LOW-PASS BANDLIMITED NOISE.

A1.1 Variability of the Running Crosscorrelation Function

The running crosscorrelation is assumed to have a rectangular time-window and is defined as follows:

$$^{k}R_{xy}(\tau,T) = \frac{1}{T}\int_{0}^{T} {}^{k}x(t) \cdot {}^{k}y(t+\tau)\,dt \qquad \ldots \quad (A1)$$

where $x(t)$, $y(t)$ are the signals at each input.

prefix k indicates the ensemble member.

$^{k}R_{xy}(\tau,T)$ varies with time. It is desired to obtain the variance of this function. By definition,

$$\mathrm{Var}\left[{}^{k}R_{xy}(\tau,T)\right] = \overline{{}^{k}R^{2}_{xy}(\tau,T)} - R^{2}_{xy}(\tau) \qquad \ldots \quad (A2)$$

$$\text{where } R_{xy}(\tau) = \overline{{}^{k}R_{xy}(\tau,T)}$$

If the signals at the two inputs are identical save for a time shift δT which is considerably less than the averaging time T, then the crosscorrelation reduces to autocorrelation with the peak of the autocorrelation function shifted by an amount δT on the τ-axis. The problem therefore reduces to finding

$$\mathrm{Var}\left[{}^{k}R_{xx}(\tau,T)\right] = \overline{{}^{k}R^{2}_{xx}(\tau,T)} - R^{2}_{xx}(\tau) \qquad \ldots \quad (A3)$$

For normal random noise (A3) reduces to, (Bendat 1958, p.259);

$$\mathrm{Var}\left[{}^{k}R_{nn}(\tau,T)\right] = \frac{2}{T^2}\int_0^T (T-\nu)\left[R_{nn}^2(\nu)+R_{nn}(\tau+\nu)R_{nn}(\tau-\nu)\right] d\nu \qquad \ldots \quad \text{(A4)}$$

The noise is assumed to be ideally bandlimited to $\frac{1}{2\pi}$ W c.p.s.* with average power P_n watts/unit bandwidth. Hence

$$R_{nn}(\tau) = \int_0^W P_n \cos \omega\tau \, d\omega$$

$$\text{i.e. } R_{nn}(\tau) = P_n W \frac{\sin W\tau}{W\tau} \qquad \ldots \quad \text{(A5)}$$

The desired result is obtained by substituting (A5) into (A4). The resulting integral may be evaluated graphically by summing the area under $R_{nn}^2(\nu)$ and $R_{nn}(\tau+\nu)R(\tau-\nu)$ for various values of τ. This has been done, assuming $T \gg \nu$ in the region $R_{nn}^2(\nu) \gg 0$, and the result is shown in Fig.6.4. It has been assumed that $P_n W = 1$ and $\mathrm{Var}\left[{}^{k}R_{nn}(0,T)\right] = 1$, without loss of generality.

The variance of ${}^{k}R_{nn}(\tau,T)$ is inversely proportional to both W and T, provided T is large. This is obvious from equations (A4) and (A5). At $\tau = 0$,

$$\mathrm{Var}\left[{}^{k}R_{nn}(0,T)\right] \doteqdot \frac{2\pi}{WT} \qquad \ldots \quad \text{(A6)}$$

for T large.

---

* i.e. the power spectrum is flat in the range 0 to $\frac{1}{2\pi}$ W c.p.s. with zero power outside this band.

A1.2 Fitting the Model to Experimental Data

In Fig.A.1 the region of the correlation function above AB is defined as the peak, where c = the critical depth and d = spread of peak. A peak in the correlation pattern corresponds to a sound image. If a single noise source with interaural delay δT is used the peak should occur at $\tau = \delta T$; the height of the peak and adjacent regions of the curve will fluctuate about their mean values (see Fig.6.4). Fluctuations in adjacent regions of the curve are highly correlated, and if the critical depth c is small enough the fluctuations at A and B should be almost identical to those at X (Fig.A.1).

The variability of a listener's lateral judgements is defined as

$$\left(\frac{1}{\beta}\right)^2 = \sigma^2 = \sigma_s^{\,2} + \sigma_c^{\,2} \qquad \ldots \qquad \text{(A7)}$$

where $\sigma_s^{\,2}$ = component due to factors characteristic of the subject, including random variations in medial plane position

$\sigma_c^{\,2}$ = component due to the limited resolution of the peak of the correlation pattern.

The results of experiments B4 and B5 have indicated that $\sigma_s \gg \sigma_c$ for low-pass bandlimited noise with a cut-

FIG. A1 : CORRELATION PEAK FOR LOW-PASS BANDLIMITED NOISE

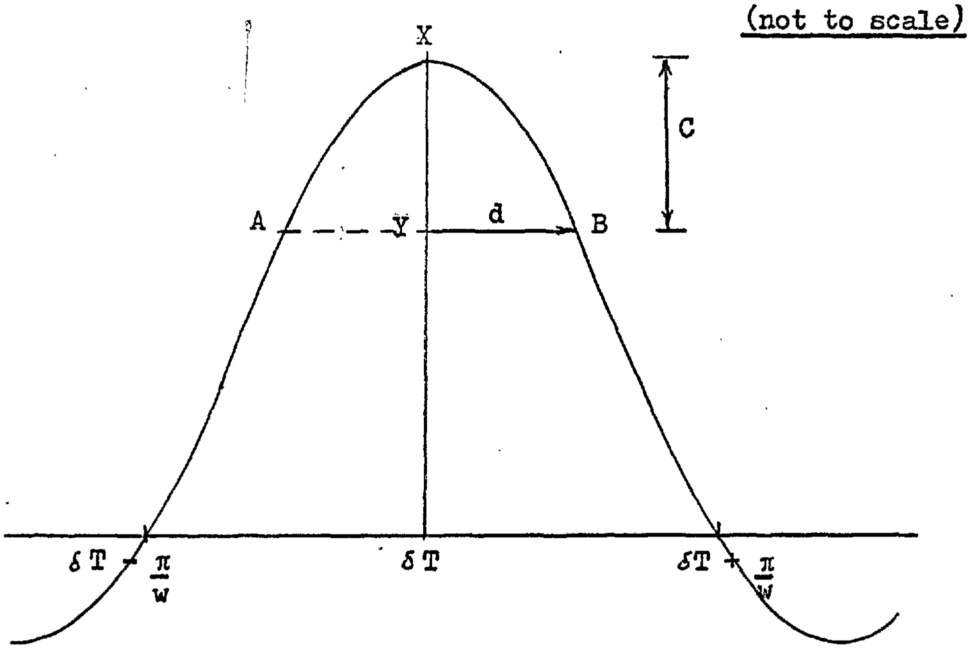


TABLE A1 : ESTIMATED PARAMETERS OF MODEL

Parameter	σ_s^2	σ_c^2	k_s	k_c
Subject R.W.	4.5	negligible	4.59	5000
" L.M.	8.0	"	9.72	"
" I.C.	9.9	"	6.37	"
" E.W.	9.8	"	10.69	"

Mean square deviation of observed data from predicted values $= 2.96$ units^2

Mean sampling variance of observations $= 3.76$ units^2

1 unit = 5.8 microsecs.

off frequency in the range 900 to 4560 c.p.s. The values of σ_s^2 have been estimated directly from the data (Table A1).

If two sources with similar values of δT are used the variability of the correlation pattern will blur the dip between the two expected peaks and a single peak will be apparent, but with an effective spread greater than that for either of the original peaks. A peak of this type is referred to as a <u>dual</u> peak. The spread of a peak fluctuates with time owing to the variability of the pattern. The ability of a listener to distinguish between a single and dual peak is defined as follows:

$$DL\,(\delta T_1 - \delta T_2) = k_s + k_c' \quad \text{time units} \qquad \ldots \qquad \text{(A8)}$$

where $DL(\delta T_1 - \delta T_2)$ = that critical spacing on the δT axis where the listener detects a dual peak 50% of the time.

k_s = a constant characteristic of the subject.

k_c' = a constant characteristic of the correlation pattern, and is equal to that critical spacing when the spread of a dual peak is proportionately greater than that of a single peak by a fixed amount.

To a first approximation assuming c small, k_c' is inversely proportional to bandwidth and hence (A8) reduces to

$$DL\ (\delta T_1 - \delta T_2) \;=\; k_s + \frac{1}{W}k_c \qquad \ldots \qquad (A9)$$

where $k_c = Wk_c'$ and is independent of bandwidth.

Using the data of experiment B3 estimates of k_s and k_c have been obtained (Table A1). The mean square deviation of the observed data from predictions of the model was found to be 2.96 units^2, with 15 degrees of freedom. The estimated sampling variance of the estimates was 3.76 units^2, thus indicating a satisfactory fit.

APPENDIX A2 A SIMPLE METHOD OF PROGRAMMING A DIGITAL COMPUTER FOR ANALYSIS-OF-VARIANCE COMPUTATIONS

A2.1 Introduction

Analysis-of-variance calculations are extremely common in statistical analysis and several methods of simplifying the computations have been developed (e.g. Ayers and Stanley 1952, Bainbridge, Grant and Radok 1956, Edwards and Horst 1950, Hartley 1956). It is important, however, to distinguish between methods designed for desk calculators and those designed specifically for digital computers. In the former case a reduction in the complexity of the arithmetic is often desired; in the latter case complexity of the arithmetic may be of little consequence provided simple rules for the algebra can be found. Furthermore, if simple patterns in the algebra exist then the computer can be programmed to generate the necessary algebraic expressions, as well as evaluate them numerically. Analysis-of-variance calculations are usually very similar in form and hence amenable to the above approach. In particular, the expressions obtained in analysing a factorial design show a remarkable symmetry.

The approach suggested in this note is based on a sorting procedure using a binary representation for the various terms occurring in the analysis. The advantages of this method are briefly;

a) patterns in the sorting procedure are easily recognized and can be exploited to provide a simple, compact programme, and

b) it is possible to feed in with the data coded expressions for any desired sum of squares. This may seem an unusual approach, but considering the complexity of a generalised analysis-of-variance programme it has distinct practical advantages: a short basic programme and a set of special data "headings", (i.e. coded programming instructions for a given experimental design) are all that is required for an entire library of statistical routines. Furthermore, this technique involves less computer time and/or memory and is extremely flexible; new designs may be catered for or special sums of squares evaluated as required.

A2.2 Definitions

A2.2.1 Nomenclature

Given factors A, B, C, D, ... each factor having a total number of levels I, J, K, L ... respectively, then $X_{ijkl\ldots}$ represents the observation corresponding to factor A having level i

" B " " j

" C " " k

factor D having level 1

and so on.

A dot in place of a suffix indicates that an arithmetic average has been taken over all values of the factor concerned.

For example $X_{i..} = \frac{1}{JK} \sum_{j,k} X_{ijk}$

A2.2.2 Classification of Averages

The averages are classified into groups according to the number of factors which are necessary to specify the average. For example, in a 3-factor design,

$X_{...}$ = zero-order average (or grand mean)

$X_{i..}$ $X_{.j.}$ $X_{..k}$ are 1st-order averages

$X_{ij.}$ $X_{i.k}$ $X_{.jk}$ are 2nd-order averages

X_{ijk} is a 3rd-order average.

The averages are also classified into types. Each type of average has associated with it a binary number which is obtained by replacing the dots in the suffix by zeros and the symbols by ones. Each digit in the binary number is said to occupy a column. Table A2.1 shows the averages of a 3-factor design arranged in ascending order of type number.

The set of averages of a given type may be classi-

TABLE A2.1

Classification of Averages (3 factors)

Type of Average	Associated Binary Number	Type Number	Order of Average
$X_{...}$	0 0 0	0	0'th
$X_{i..}$	1 0 0	1	1st
$X_{.j.}$	0 1 0	2	1st
$X_{ij.}$	1 1 0	3	2nd
$X_{..k}$	0 0 1	4	1st
$X_{i.k}$	1 0 1	5	2nd
$X_{.jk}$	0 1 1	6	2nd
X_{ijk}	1 1 1	7	3rd

TABLE A2.2

Level Count for a Type 5 Average

(i.e. $X_{i.k}$ where I and K are 3&2 respectively)

Symbol	Level Count	Level Number
$X_{1.1}$	1 . 1	1
$X_{2.1}$	2 . 1	2
$X_{3.1}$	3 . 1	3
$X_{1.2}$	1 . 2	4
$X_{2.2}$	2 . 2	5
$X_{3.2}$	3 . 2	6

fied in ascending order according to the following counting procedure: all suffixes in the average type are set to unity. The first suffix (going from left to right) is increased in steps of unity until its maximum value is reached. The second suffix is then increased by unity and the process is repeated until the second suffix reaches its maximum value. The third suffix is then increased by unity; and so on until all the suffixes reach their respective maximum values. This process is termed a level count and there is a corresponding level number for each stage of the process. Table A2.2 shows the level count for an average of type 5.

If in the above counting procedure the maximum number of levels for each digit is fixed at two then it is a binary count. If the count begins at the highest possible value and counts downwards, it is referred to as a decount.

The values of the various averages are stored in registers. The registers are numbered consecutively, where each number is the address of the associated register.

A2.3 Description of Basic Programme

The data are read in and all relevant averages are computed and stored in one long column of registers.

These averages are stored in descending order of type number, but within each type in ascending order of level number. Table A2.3 shows how the averages for a simple 3-factor design are stored. This method of computing and storing the averages may be conveniently carried out by repetitive counting operations (see Diagram A1).

The general address for any average is given by

$$A_{ijk---} = A_{111---} + (i-1) + I(j-1) + IJ(k-1) + \ldots$$

where A_{111---} = address of register storing the first average of the type under consideration.

Hence the address of any desired average is obtained directly from its type specification.

The second stage of the programme (Diagram A2) selects those averages which are required in evaluating a given sum of squares. The task is essentially one of sorting according to a set of rules and two approaches are possible. One method is to feed in the appropriate formulae, suitably coded, with the data. For example, the sum of squares

$$\sum_{i,k} \left[X_{i.k} - X_{i..} - X_{..k} + X_{...}\right]^2 \quad \text{can be}$$

represented by 5 - 1 - 4 + 0, where the numbers refer to the type of average (see Table A2.1). The above

TABLE A2.3

Storing Averages Obtained in a 3-Factor Design

(I, J and K = 2, 3 and 2 respectively)

Register Address	Average	Type Number	Level Number	Register Address	Average	Type Number	Level Number
1	X_{111}	7	1	19	$X_{1.1}$	5	1
2	X_{211}	7	2	20	$X_{2.1}$	5	2
3	X_{121}	7	3	21	$X_{1.2}$	5	3
4	X_{221}	7	4	22	$X_{2.2}$	5	4
5	X_{131}	7	5	23	$X_{..1}$	4	1
6	X_{231}	7	6	24	$X_{..2}$	4	2
7	X_{112}	7	7	25	$X_{11.}$	3	1
8	X_{212}	7	8	26	$X_{21.}$	3	2
9	X_{122}	7	9	27	$X_{12.}$	3	3
10	X_{222}	7	10	28	$X_{22.}$	3	4
11	X_{132}	7	11	29	$X_{13.}$	3	5
12	X_{232}	7	12	30	$X_{23.}$	3	6
13	$X_{.11}$	6	1	31	$X_{.1.}$	2	1
14	$X_{.21}$	6	2	32	$X_{.2.}$	2	2
15	$X_{.31}$	6	3	33	$X_{.3.}$	2	3
16	$X_{.12}$	6	4	34	$X_{1..}$	1	1
17	$X_{.22}$	6	5	35	$X_{2..}$	1	2
18	$X_{.32}$	6	6	36	$X_{...}$	0	1

combination of average types is then evaluated for each permutation of factor levels and the sum of squares over all relevant factors is obtained. Since the above, coded expression involves integers only, it is possible to,

a) feed in expressions of this type with the data;

b) use decimal fractions for special coding purposes without affecting the meaning of the expression, (.e.g such as indicating relative weighting in the summing process).

On the other hand, if a large number of sums of squares is required it may be convenient to generate the required expressions within the computer. The process of generating the necessary formulae is considerably simplified if the design being analysed displays some symmetry. For example, in a factorial design all the standard sums of squares may be obtained from the following general expression:

$$\begin{array}{l}\text{Type k}\\ \text{sum of}\\ \text{squares}\end{array} = \sum \Bigg[(-1)^n \begin{bmatrix}\text{Average of}\\ \text{type k}\\ \text{order n}\end{bmatrix} + (-1)^{n-1} \begin{bmatrix}\text{Sum of all averages}\\ \text{of order(n-1) which}\\ \text{can be derived from}\\ \text{previous average of}\\ \text{order n}\end{bmatrix} + (-1)^{n-2} \begin{bmatrix}\text{Sum of all averages}\\ \text{of order(n-2) which}\\ \text{can be derived from}\\ \text{previous averages of}\\ \text{order (n-1)}\end{bmatrix} + \text{-----------------} + (-1)^{o} \begin{bmatrix}\text{zero-order}\\ \text{average}\end{bmatrix} \Bigg]^2$$

For illustration, the various sums of squares for a 3-level factorial design are given in Table A2.4. Note that these sums of squares are classified according to the type number of the highest order average occurring in the expression. Note further that a simple rule for determining whether or not an average of a given type can be derived from averages of the next highest type, is obtained by comparing their binary representations; if the lower order average has a one in any column where the higher order average has a zero, then it is not possible to derive the one type of average from the other. For example, in the sum-of-squares expression number 5 (Table A2.4) a second-order average of the form $X_{.j.}$ does not occur since $X_{.j.}$ cannot be derived from averages of the form $X_{i.k}$. The corresponding binary numbers are 010 and 101 respectively. Diagram A3 shows the sorting routine for a factorial design as used in a programme designed for MERCURY.

A2.4 Summary

The programme consists of two parts;

a) a basic routine which computes and stores all averages which are relevant to a given analysis of variance,

TABLE A2.4

Component Sums of Squares in a 3-Factor Design

Source	Sum of Squares	Binary Representation	Average of highest type occurring
A effect	$JK\sum_i \left[X_{i..} - X_{...}\right]^2$	100 - 000	1
B effect	$IK\sum_j \left[X_{.j.} - X_{...}\right]^2$	010 - 000	2
AB interaction	$K\sum_{ij} \left[X_{ij.} - X_{i..} - X_{.j.} + X_{...}\right]^2$	110- 100+010 + 000	3
C effect	$IJ\sum_k \left[X_{..k} - X_{...}\right]^2$	001 - 000	4
AC interaction	$J\sum_{ik} \left[X_{i.k} - X_{i..} - X_{..k} + X_{...}\right]^2$	101 - 100+001 + 000	5
BC interaction	$I\sum_{jk} \left[X_{.jk} - X_{.j.} - X_{..k} + X_{...}\right]^2$	011 - 010+001 + 000	6
ABC interaction	$\sum_{ijk} \left[X_{ijk} - X_{ij.} - X_{i.k} - X_{.jk} + X_{i..} + X_{.j.} + X_{..k} - X_{...}\right]^2$	111 - 110+101+011 + 100+010+001 - 000	7

b) a sub-routine which when coupled to (a) computes all main-effect and interaction sums of squares for a factorial design. The only limitation in using this routine is that the design must be complete with an equal number of replications per cell. There is no limit to the number of factors or their respective levels, although in practice the size of the computer may limit the total number of averages which can be handled.

In place of (b) an appropriate sub-routine for another type of analysis may be used. If desired, the programming instructions for (b) may be conveniently fed in with the data. An important feature of the programme is its highly repetitive nature; the major subsections are all variations of the simple counting procedure described in section A2.2.2. The end result is a neat and compact programme (see Diagram A4).

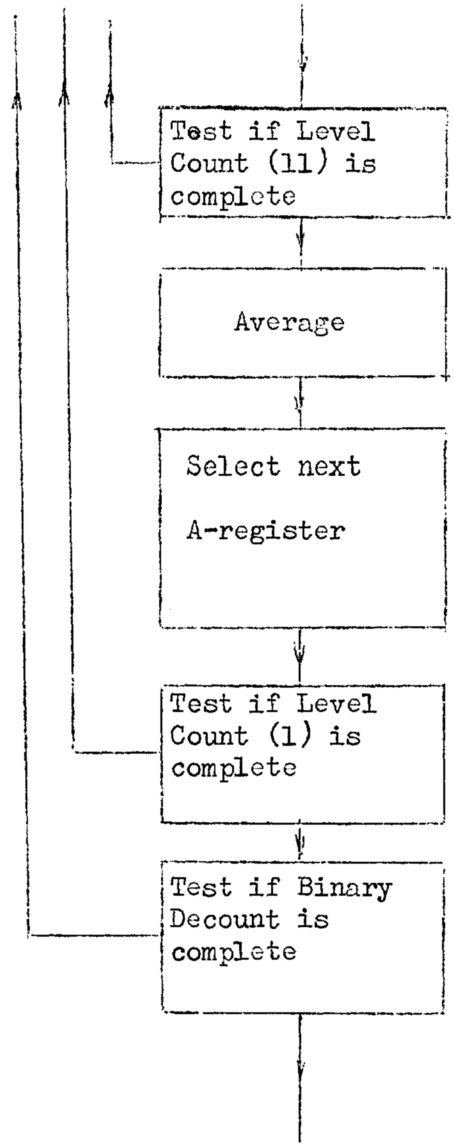
The author has recently observed that a programme similar in some respects to that described above, but of wider general applicability, has been reported by Yates, Gower and Simpson (1963). The price of generality, however, is a considerably more complex programme.

DIAGRAM A1 : COMPUTING AND STORING AVERAGES

Stage No.	Flow Diagram	Description
1	Clear Registers Set Counting Routine	If there are N factors then two sets of N registers are set aside for doing binary and level counts with N digits.
2	Read Data	Data read in and stored in ascending level number starting at register A_1
3	Binary Decount	This enumerates in descending order all the types of averages to be computed. Note that the highest type of average = original data and these have been read in directly in their correct order.
4	Level Count (1)	Enumerates in ascending level number all averages of the type specified by stage 3. Note: if any digit of the binary number is zero then the corresponding factor is omitted in this level count.
5	Level Count (11)	Enumerates (in ascending level number) all the individual terms over which the average given by stages 3 and 4 is taken. Count procedure as in 4, but in this case a factor is omitted if corresponding binary digit is unity.
6	Address Location	Locates address of the register storing the term given by 5 above. A general expression for the address of any average is given in section A2.3.
7	Summation	Adds the contents of the register located by 6 to the contents of the register reserved for storing the average currently given by stages 3 and 4.

... continued on next page ...

DIAGRAM A1 (cont'd.)



8 If the count is incomplete increase Level Count (11) by unity and repeat 6,7,8; otherwise go to 9.

9 Divide the sum found in 7 by the total number of levels in 5. (Stages 7 and 9 represent the averaging process.)

10 Select the next register in the column for storing the forthcoming average. The counting procedure in 3 and 4 ensures that averages are ordered in descending order of type number and ascending order of level number.

11 Similar to 8

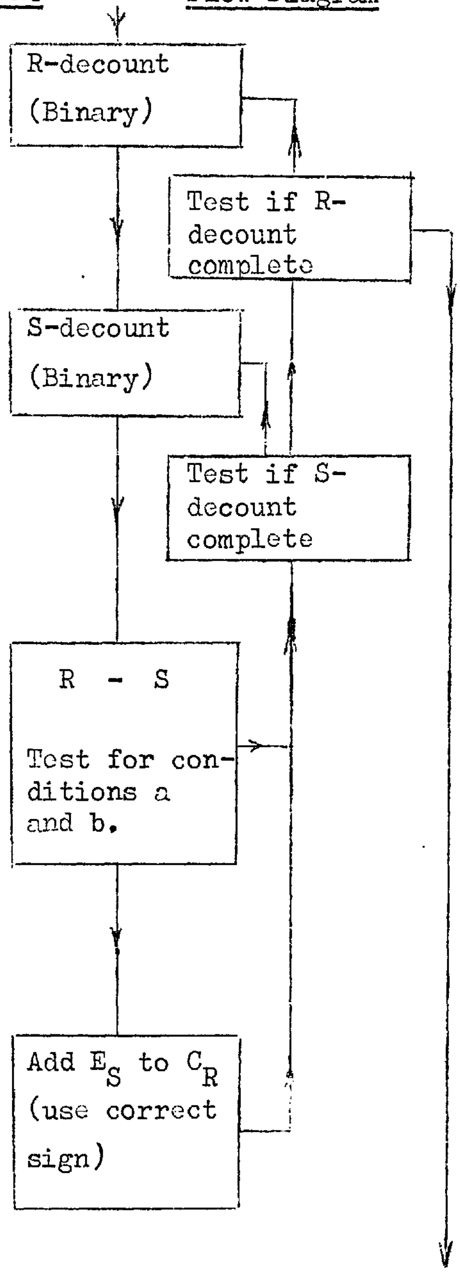
12 Similar to 8

DIAGRAM A2 : SUMMING SQUARES

Stage No.	Flow Diagram	Description
13	Level Count	The count is over all levels of all factors. (The contributions to the various sums-of-squares are found separately for each permutation of factor levels.)
14	Binary Decount	For each permutation of factor levels given by 13 the various types of average are located and stored in a special set of registers. (E_1 to E_R say).
15	Address Location	
16	Store Corresponding Average	
17	Test if Binary Decount is complete	
18	Sorting Routine	The various averages obtained in stages 14 to 17 are sorted either to a programmed set of rules (see diag.A3) or according to information fed in with the data. The several combinations of averages thus derived are stored in a set of registers, (C_1 to C_q, say).
19	Square and Add	The contents of each C-register is squared and added to the contents of a corresponding set of D-registers, (D_1 to D_q, say).
20	Test if Level Count is complete	
21	Print Out	Print out contents of D-registers. The D-registers contain the required component sums-of-squares.

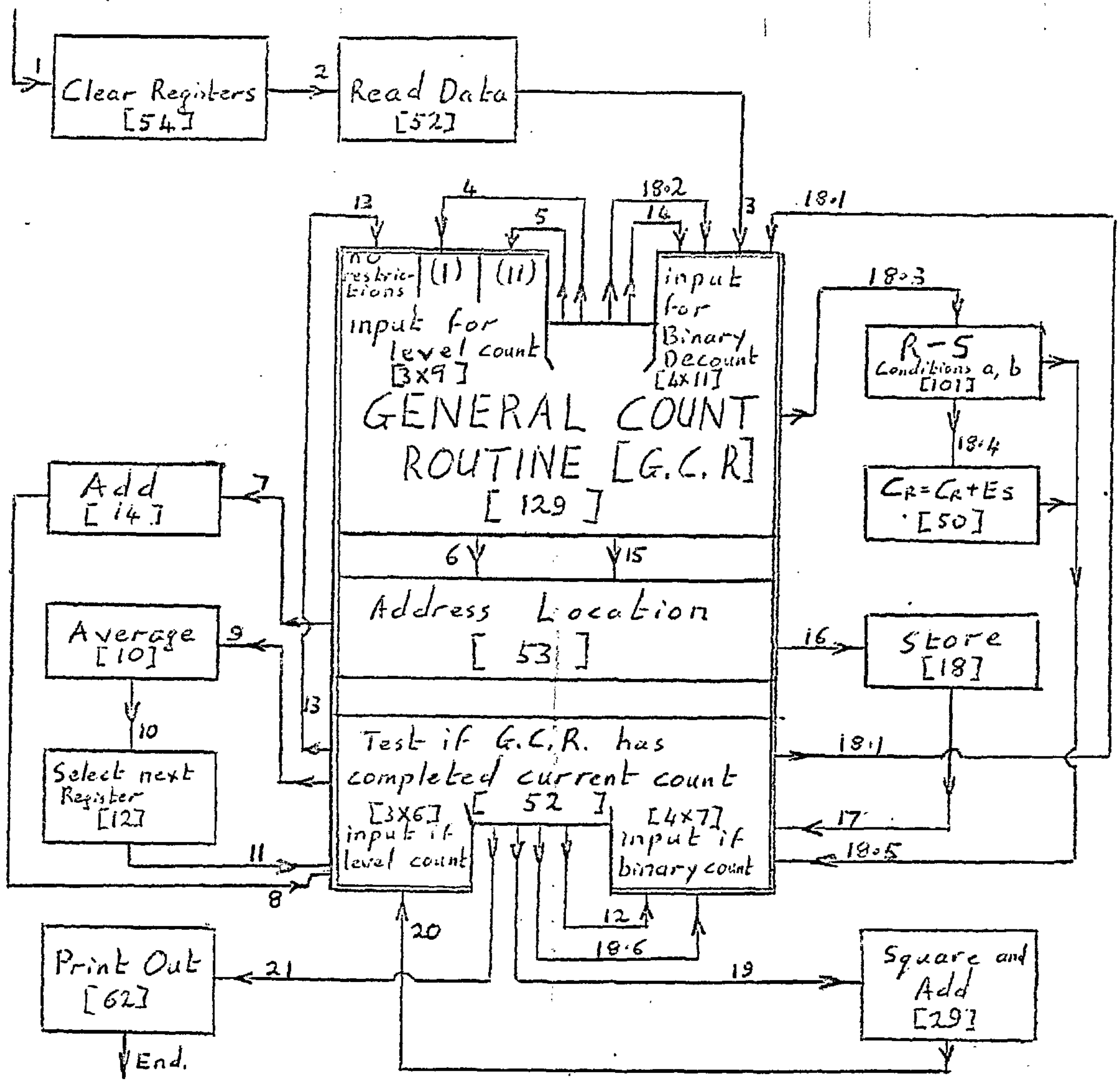
DIAGRAM A3 : SORTING ROUTINE FOR A FACTORIAL DESIGN

Flow Diagram



Stage No	Description
18.1	Count begins at R = max. type number and descends to zero.
18.6	If R-decount is complete go on to next routine otherwise continue with R-decount.
18.2	Count begins at R-1 and descends to zero.
18.5	If S-decount is complete go on to R-decount, otherwise continue with S-decount.
18.3	Compare the binary numbers S and R. Condition a: R and S differ in only one digit where that digit is a zero in S. Condition b: The current value of S has not already satisfied condition a for a previous value of R.
18.4	The average corresponding to binary number S (which has been stored in register E_S (stage 16)) is added to contents of register C_R. If odd number of zeros in S use negative sign.

DIAGRAM 4. Compact Form of Flow Diagram for a Complete Factorial Design.



Notes: 1) The number outside the boxes indicates the stage number corresponding to diagrams 1, 2 & 3.

2) The number in brackets within each box indicates the number of machine instructions used in a program designed for Mercury. These figures are larger than the minimum possible [by about 30%] since several tricks were employed to conserve working space at the expense of program space.

LIST OF REFERENCES

ARNOULT, M.D. (1952) Localization of sound during rotation of the visual environment. Amer. J.Psychol., 65, 48-58.

AYERS, J.D. and STANLEY, J.P. (1952) The rolling totals method of computing sums, sums of squares, and sums of cross-products. Psychometrika, 17, p 305-310.

BAINBRIDGE, J.R., GRANT, A.M. and RADOK, U. (1956), Tabular analysis of factorial experiments and the use of punch cards. J.Amer.Statist.Ass., 51, 149-158.

BANISTER, H. (1927) Auditory Theory: A criticism of Professor Boring's hypothesis. Amer.J. Psychol. 38, 436-440.

VON BEKESY, G. (1930) Zur Therie des Horens, Physikal Zsch., 31, 857-858. See also Chapter 8, Experiments in Hearing - G.vonBekesy, McGraw Hill, 1960.

VON BEKESY, G. (1948) Vibration of the head in a sound field and its role in hearing by bone conduction. J. Acoust. Soc.Amer., 20, 749-760.

VON BEKESY, G. (1956) Current status of theories of hearing. Science, 123, 779-783.

BENDAT, J.S. (1958) Principles and Applications of Random Noise Theory. John Wiley & Sons, New York.

VAN BERGEIJK, W.A. (1962) Variation on a theme of Bekesy: a model of binaural interaction. J. Acoust.Soc.Amer., 34 (9,Pt.11), 1431-1437.

BERKSON, J. (1951) Why I prefer logits to probits, Biometrics, 7, 327-339.

BERKSON, J. (1955) Maximum likelihood and minimum chi-square estimates of the logistic function. J.Amer.Statist.Assoc., 50, 130-162.

BLACKWELL, H.R. (1953) Evaluation of the neural quantum theory in vision. Amer.J.Psychol., 66, 397-408.

BLODGETT, H.C., WILBANKS, W.A. and JEFFRESS, L.A. (1956) Effect of large inter-aural time differences upon the judgement of sidedness. J. Acoust. Soc.Amer., 28, 639-643.

BOCCA, E. (1961) Factors influencing binaural integration of periodically switched signals (messages). Acta Oto-Laryngol., 53, 142-144.

BORING, E.G. (1942) Sensation and perception in the history of experimental psychology. D.Appleton-Century Company, New York 1942.

BOWLES, J.A. (1960) Binaural discriminatory processes. M.Sc.Thesis, University of London.

BROADBENT, D.E. (1955) A note on binaural fusion. Quart.J. exp.Psychol., 7, 46-47.

BROADBENT, D.E. and LADEFOGED, P. (1957) On the fusion of sounds reaching different sense organs. Acoust.Soc.Amer. 29, 708-710.

BROADBENT, D.E. (1958) Perception and Communication. Pergamon Press, London.

BROCKELSBY, C.F., PALFREEMAN, J.S., GIBSON, R.W. (1963) Ultrasonic Delay Lines. Iliffe Books Ltd., London.

BROWNLEE, K.A., HODGES, J.L. Jr. and ROSENBLATT, M. The up-and-down method with small samples. J.Amer.Statist. Ass. 1953, 48, 262-276.

BUTLER, R.A. and NAUNTON, R.F. (1962) Some effects of unilateral masking upon the localization of sound in space. J.Acoust.Soc.Amer. 34 (8), 1100-1107.

CHERRY, E.C. (1953) Some experiments on the recognition of speech, with one and with two ears. J. Acoust.Soc.Amer. 25, 975-979.

CHERRY, E.C. (1955) Information, feedback, and the human senses. Trans.Soc.Instr.Tech., 7, 81-89.

CHERRY, E.C. (Ed.). (1960). Information Theory. Butterworths, London.

CHERRY, E.C. (1961) Whey we have two ears. Annual Meeting of British Association for the Advancement of Science, Norwich, Sept.1961, reprinted in Advancement of Science, Sept.1962, 218-221.

CHERRY, E.C. (1962) The cocktail party problem. Discovery, March 1962, 32-35.

CHERRY, E.C. (1963) Some simple experiments in binaural listening. Meeting of the Institute of Physics and the Physical Society. Dept. of Elect. Eng., Imperial College of Science and Technology 13th June 1963.

CHERRY, E.C., and BOWLES, J.A. (1960) Contribution to a study of the "cocktail party problem". J.Acoust.Soc.Amer., 32, 884.

CHERRY, E.C. and SAYERS, B.Mc A. (1956). Human 'Cross-correlator'- A technique for measuring certain parameters of speech perception. J.Acoust. Soc.Amer. 28, 889-894.

CHERRY, E.C., and SAYERS, B. Mc A. (1959) On the mechanism of binaural fusion. J.Acoust.Soc.Amer., 31, 535.

CHERRY, E.C. and TAYLOR, W.K. (1954) Some further experiments upon the recognition of speech with one and with two ears. J.Acoust.Soc.Amer., 26, 554-559.

CLARK, B. and GRAYBIEL, A. (1949) The effect of angular acceleration on sound localization: the audiogyral illusion. J. Psychol., 28, 235-244.

CORLISS, Edith (1962) Personal communication.

CORNSWEET, T.N. (1962) The staircase method in psychophysics. Amer.J.Psychol. 75 (3), 485-491.

CRAMER, E.M. and HUGGINS, W.H. (1958) Creation of pitch through binaural interaction. J.Acoust. Soc.Amer., 30, 413-417.

DAVID, E.E. Jr., (1959) Comment on the Precedence Effect. 3rd Int.Cong. Acoust., Stuttgart 1959.

DAVID, E.E. (1962) Closing the binaural gap. 63rd Meeting of Acoustical Society of America, New York, May 1962.

DAVID, E.E.Jr. and HANSON, R.L. (1962) Binaural hearing and free field effects. IVth International Congress on Acoustics, Copenhagen, August 1962.

DAVID, E.E.Jr., GUTTMAN, N. and VAN BERGEIJK, W.A. (1958). On the mechanism of binaural fusion. J.Acoust. Soc.Amer., 30, 801.

DAVID, E.E.Jr., GUTTMAN, N. and VAN BERGEIJK, W.A. (1959) Binaural interaction of high-frequency complex stimuli. J.Acoust.Soc.Amer., 31, 774-782.

DAVIES, O.L. (1956) The design and analysis of industrial experiments. Oliver and Boyd, London.

DEATHERAGE, B.H. (1961) Binaural interaction of clicks of different frequency content. J.Acoust. Soc.Amer., 33, 139-145.

DEATHERAGE, B.H. and HIRSH, I.J. (1959) Auditory localization of clicks. J.Acoust.Soc.Amer., 31, 486-492.

DIXON, W.I. and MOOD, A.M. (1948) A method for obtaining and analysing sensitivity data. J.Amer. Statist.Assoc.,43, 109-126.

EDWARDS, A.S. (1955) Accuracy of auditory depth perception. J. gen.Psychol., 52, 327-329.

EDWARDS, A.L. and HORST, P. (1950). The calculation of sums of squares for interactions in the analysis of variance. Psychometrika, 15,17-24.

EGAN, J.P. (1948) The effect of noise in one ear upon the loudness of speech in the other. J.Acoust. Soc.Amer., 20, 58-62.

EGAN, J.P. (1954) Hearing. Annual Review of Psychology, 5, 89-110.

EGAN, J.P., CARTERETTE, E.C. and THWING, E.J. (1954) Some factors affecting multi-channel listening. J.Acoust.Soc.Amer., 26, 774-

EISENHART, C. (1947) Inverse sine transformation of proportions, Chapter 16 in Eisenhart, Hastay and Wallis (1947).

EISENHART, C., HASTAY and WALLIS (1947) Techniques of Statistical Analysis. McGraw Hill, New York.

FAY, R.D. and HALL, W.M. (1956) Historical note on the Haas effect. J.Acoust.Soc.Amer. 28, 131-132.

FINNEY, D.J. (1944) The application of probit analysis to the results of mental tests. Psychometrica, 9, 31-39.

FINNEY, D.J. (1947) Probit Analysis. Cambridge University Press, 1st edition 1947, or 2nd edition 1952.

FLANAGAN, J.L., DAVID, E.E.Jr. and WATSON, B.J. (1962) Physiological correlates of binaural lateralization. IVth International Congress on Acoustics, Copenhagen, August 1962.

FLETCHER, H. (1953) Speech and Hearing in Communication. D. Van Nostrand Company, Princeton, New Jersey.

FOOT, G. (1963) A preliminary investigation into the limits of binaural fusion. Special Project Report, Dept. of Elect.Eng., Imperial College of Science and Technology, London.

FOURCIN, A.J. (1961) An aspect of the perception of speech. Proceedings 4th International Congress of Phonetic Sciences, Helsinki 1961. Mouton and Co., The Hague, 1962.

FOURCIN, A.J. (1964) Private communication.

FRANSSEN, N.V. (1960) Some considerations on the mechanism of directional hearing. Thesis, Technische Hogeschool Delft, Holland.

FRIEDMAN, M. (1947) Planning an experiment for estimating the mean and standard deviation of a normal distribution from observations on the cumulative normal. Chapter 11 in Eisenhart, Hastay and Wallis (1947).

GARWOOD, F. (1941) The application of maximum likelihood to dosage mortality curves. Biometrika, 32, 46-58.

GRAYBIEL, A. and NIVEN, J.L. (1951) The effect of a change in direction of resultant force on sound localization: the audiogravic illusion. J. exp. Psychol., 42, 227-230.

GUILFORD, J.P. (1954) Psychometric methods, 2nd edition. McGraw-hill Book Company, New York.

GUTTMAN, N. (1962) A mapping of binaural click lateralizations. J. Acoust.Soc.Amer., 34, 87-91.

GUTTMAN, N., VAN BERGEIJK, W.A. and DAVID, E.E.Jr. (1960). Monaural temporal masking investigated by binaural interaction. J.Acoust.Soc.Amer., 32, 1329-1336.

HARRIS, G.G. (1960) Binaural interaction of impulsive stimuli and pure tones. J.Acoust.Soc.Amer. 32 (6), 685-692.

HARRIS, G.H., FLANAGAN, J.L. and WATSON, B.J. (1963). Binaural interaction of a click with a click pair. J.Acoust.Soc.Amer., 35, 672-678.

HARTLEY, H.O. (1956) A plan for programming analysis of variance for general purpose computers. Biometrics, 12, 110-122.

HELD, R. (1955) Shifts in binaural localization after prolonged exposure to atypical combinations of stimuli. Amer.J.Psychol., 68, 526-548.

HIRSH, I.J. (1948). Binaural summation - a century of investigation. Psychol. Bull., 1948, 45, 193-206.

HIRSH, I.J. (1950) Relation between localization and intelligibility. J.Acoust.Soc.Amer., 22, 196 -

HIRSH, I.J. and POLLACK I. (1948) The role of interaural phase in loudness. J.Acoust.Soc.Amer., 20, 761-66.

HOLMES. J.N. and DUKES, J.M.C. (1954) A speech-waveform correlator with magnetic-tape delay and electronic multiplication. Proc.I.E.E., 101, Pt.III, 225-237.

INGHAM, J.G. (1959) Variations in the cross-masking with frequency. J. exp. Psychol., 1959 (Sep), 58, 199-205.

IRWIN, J.O. (1937) Statistical method applied to biological assays. J. Roy. Statist.Soc., 4, 1.

IRWIN, J.O. and CHEESEMAN, E.A. (1939) On the maximum-likelihood method of determining dosage-response curves and approximations to the median-effective dose, in cases of a quantal response. J.Roy.Statist.Soc. Supplement, Vol. VI, No.2, 174.

JACKSON, C.V. (1953) Visual factors in auditory localization. Quart.J. exp.Psychol., 5, 52-65.

JEFFRESS, L.A. (1948) A place theory of sound localization. J. Comp. Physiol.Psychol., 41, 35-39.

JEFFRESS, L.A., BLODGETT, H.C.(1962a)* Effect of interaural correlation on the precision of centering a noise. J. Acoust.Soc.Amer., 34 (8), 1122-1123.

JEFFRESS, L.A. and BLODGETT, H.C. (1962b) Effect of switching earphone channels upon the precision of centering. J. Acoust. Soc. Amer., 34, 1275-1276

JEFFRESS, L.A., BLODGETT, H.C., SANDEL, T.T. and WOOD,C.L. (1956) Masking of tonal signals. J.Acoust. Soc.Amer., 28, 416-426.

JEFFRESS, L.A. and TAYLOR, R.W. (1961) Lateralization vs. localization. J.Acoust.Soc.Amer., 33, 482-483.

JONGKEES, L. and VAN DER VEER, R. (1958) On directional sound localization in unilateral deafness and its explanation. Acta Oto-Laryngol., 49, 119- 131.

KIKUCHI, Y. (1957) Objective allocation of sound-image from binaural stimulation. J. Acoust.Soc. Amer., 29, 124-128.

KITAGAWA, S. and SHINTAKU, Y. (1957) Direction perception and subjective auditory direction line. Acta Oto-Laryngol., 47, 431-443.

KLENSCH, H. (1948) Beitrag zur Frage der Lokalisation des Schalles im Raum. Pflüg.Arch.ges.Physiol. 250, 492-500. Abstracted in Psychological Abstracts, 1950, No.4449.

KLUMPP, R.G. and EADY, H.R. (1956) Some measurements of interaural time difference thresholds. J.Acoust.Soc.Amer., 28, 859-860.

* Should read JEFFRESS, L.A., BLODGETT, H.C. and DEATHERAGE, B.H. (1962)

KOCK, W.E. (1950) Binaural localization and masking. J.Acoust.Soc.Amer., 22, 801-804.

KRAUSKOPF, J. (1954) Figural after-effects in auditory space. Amer.J.Psychol., 67, 278-287.

LEAKY, D.M. (1959) Some measurements on the effects of inter-channel intensity and time differences in two channel sound systems. J.Acoust.Soc. Amer., 31, 977-986.

LEAKEY, D.M., SAYERS, B.M. and CHERRY C. (1958) Binaural fusion of low- and high-frequency sounds. J.Acoust.Soc.Amer., 30, 222-223.

LEVITT, H. (1960) A novel method of analysing simple, recurrent, reactive circuits. Internal Note, Dept. Elect.Eng., Imperial College of Science and Technology, London.

LEVITT, H. (1961) An extension of Pascal's triangle. Math.Gaz., 45, 99-107.

LEVITT, H. (1962) Discussion on "A method of calculating the transfer functions of ladder networks". Proc.I.E.E., 109C, 290-291.

LICKLIDER, J.C.R. (1948) The influence of interaural phase relations upon the masking of speech by white noise. J.Acoust.Amer., 20, 150-159.

LICKLIDER, J.C.R. (1951) A duplex theory of pitch perception. Experienta, 7, 128-134.

LICKLIDER, J.C.R. (1952) On the process of speech perception. J.Acoust.Soc.Amer., 24, 590-594.

LICKLIDER, J.C.R. (1959) Three auditory theories; in S.Koch.(ed) Psychology: A Study of a science. Vol.1, (Study 1, Conceptual and systematic), 41-144. McGraw-Hill, New York.

MATHES, R.C. (1955) Monaural direction finding. J.Acoust.Soc.Amer., 27, 792.

MAXWELL, A.E. (1959) Maximum likelihood estimates of item parameters using the logistic function. Psychometrika, 24, 221-227.

McALEER, H.T. (1962) Precise delay measurement. General Radio Company Reprint No.A-94, see also Electronics, 13th January 1961.

McCARTHY, P.J. (1949) A class of method for estimating reaction to stimuli of varying severity. J.educ.Psychol., 40, 143-56.

MILLER, G.A. and GARNER, W.R. (1944) Effect of random presentation on the psychometric function: implications for a quantal theory of discrimination. Amer.J.Pscyhol., 57, 451-467.

MILLS, A.W. (1960) Lateralization of high-frequency tones. J.Acoust.Soc.Amer., 32, 132-134.

MØLLER, A.R. (1962) Acoustic Reflex in Man. J.Acoust. Soc.Amer., 34, 1524-1534.

MOOD, A.M. (1950) Introduction to the theory of statistics. McGraw-Hill Book Company, New York.

MORAY, N. (1960) Studies of Selective Listening. Ph.D.Thesis, University of Oxford.

MOUSHEGIAN, G., and JEFFRESS, L.A. (1959) Role of interaural time and intensity differences in the lateralization of low-frequency tones. J.Acoust.Soc.Amer., 31, 1441-1445.

MOUZON, J.C. (1955) Stereophonic hearing with one earphone. J.Acoust.Soc.Amer., 27, 381.

MUNSON, W.A. and WIENER, F.M. (1950) Sound measurements for psychophysical tests. J.Acoust.Soc.Amer., 22, 382-386.

MUNSON, W.A. and WIENER, F.M. (1952) In search of the missing 6 dB. J.Acoust.Soc.Amer., 24, 498-501.

NORLUND, B. (1962) Physical factors in angular localization. Acta Oto-laryngol., Stockholm, 1962 54 (1), 75-93.

PAGE, E.S. (1955) A test for a change in a parameter occurring at an unknown point. Biometrika, 42, 523-527.

PAGE, E.S. (1957) On problems in which a change in a parameter occurs at an unknown point. Biometrika, 44, 248-252.

PIKLER, A.G. and HARRIS, J.D. (1955) Channels of reception in pitch discrimination. J.Acoust. Soc.Amer., 1955, 27, 124-131.

POLLACK, I. (1959) Binaural communication systems - Preliminary examination. J.Acoust.Soc.Amer., 31, 81-82.

POLLACK, I. (1960) Temporal sampling parameters of interaural noise correlations. J.Acoust. Soc.Amer., 32, 795-799.

POLLACK, I. and PICKETT, J.M. (1958) Stereophonic listening and speech intelligibility against voice babble. J.Acoust.Soc.Amer., 30, 131.

POLLACK, I. and TRITTIPOE, W.J. (1959a) Binaural listening and interaural noise cross correlation. J.Acoust.Soc.Amer., 31, 1250-1252.

POLLACK, I. and TRITTIPOE, W.J. (1959b) Interaural noise correlations: examinations of variables. J.Acoust.Soc.Amer., 31, 1616-1618.

QUASTLER, H. (Ed.) (1955) Editor - Information Theory in Psychology. The Free Press, Illinois.

QUANDT, R.E. (1958) The estimation of the parameters of a linear regression system obeying two separate regimes. J.Amer.Statist.Assoc., 53, 873-880.

QUANDT, R.E. (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes. J.Amer.Statist.Assoc., 55, 324-330.

RAAB, D.H. and OSMAN, E. (1962) Effect of masking noise on lateralization and loudness of clicks. J.Acoust.Soc.Amer., 34 (10), 1620-1624.

REAM, N. (1961) A method of calculating the transfer functions of ladder networks. Monograph No. 432M, Institution of Electrical Engineers, also Proc.I.E.E., 108C, 354-

REAM, N. (1962) The limitations of RC ladder networks as process simulators. Proc.I.E.E., 109B, 529-536.

REYNOLDS, G.S. and STEVENS, S.S. (1960) Binaural summation of loudness. J.Acoust.Soc.Amer., 32, 1337-1344.

REYNOLDS, A.J., VOELCKER, H.B.Jr., and WHITE, R.E.C. (1961) A technique for improving the intelligibility of speech transmitted by amplitude modulation over noisy channels. I.R.E. Transactions on Communication Systesm, CS-9, 222-225.

RICHARDS, D.L. (1952) Design and analysis of subjective acoustical experiments which involve a quantal response. Acustica 2, 83-91.

ROSENBLATT, F. (1962) Application of perceptrons to speech recognition. Speech Communication Seminar, Royal Institute of Technology, Sweden. August 1962.

ROSENZWEIG, M.R. (1961) Development of research of the physiological mechanisms of auditory localization. Psychol.Bull., 58, 376-389.

RUDMOSE, W. (1962) Pressure vs free field thresholds at low frequencies. IVth International Congress on Acoustics, Copenhagen, August 1962.

SANDEL, T.T., TEAS, D.C., FEDDERSEN, W.E. and JEFFRESS, L.A. (1955) Localization of sound from single and paired sources. J.Acoust.Soc.Amer., 27, 842-852.

SAYERS, B.McA. (1957) Studies of aural perception in speech communication channels. Ph.D.Thesis, University of London.

SAYERS, B.McA. and CHERRY, E.C. (1957) Mechanism of binaural fusion in the hearing of speech. J.Acoust.Soc.Amer., 29, 973-987.

SCHUBERT, E.D. (1956) Some preliminary experiments on binaural time delay and intelligibility. J.Acoust.Soc.Amer., 28, 895-901.

SCHUBERT, E.D. (1963) Interpretation of the Butler-Naunton localization shifts. J.Acoust. Soc.Amer., 35 (1), 113.

SCHUBERT, E.D. and ELPERN, B.S. (1959) Psychophysical estimate of the velocity of the traveling wave. J.Acoust.Soc. Amer.31, 990-994.

SCHUBERT, E.D. and PARKER, C.D. (1955) Addition to Cherry's findings on switching speech between the two ears. J.Acoust.Soc.Amer., 27, 792-794.

SCHUBERT, E.D. and SCHULTZ, M.C. (1962) Some aspects of binaural signal selection. J.Acoust. Soc.Amer., 34 (6), 844-849.

SHANNON, E.C. and WEAVER, W. (1949) The mathematical theory of communication. The University of Illinois Press.

SHAW, E.A.G. (1962) On the real ear calibration of circumaural earphones. Report No.APS-417, Division of Applied Physics, National Research Council of Canada.

SMITH, M.H. and LICKLIDER, J.C.R. (1949) Statistical bias in comparisons of monaural and binaural thresholds: binaural summation or binaural supplementation. Psychol.Bull., 46, 278-284.

SNOW, W.B. (1954) Effect of arrival time on stereophonic localization. J.Acoust.Soc.Amer., 26, 1071-1074.

SPENCER, A.J. (1963) Measurement of phase angle using a counter. Marconi Instrumentation, 9, 60-63.

TEAS, D.C. (1962) Lateralization of acoustic transients. J.Acoust.Soc.Amer. 34 (9, Pt.II), 1460-1465.

THURLOW, W.R. and ELFNER, L.F. (1959) Pure-tone cross-set localization effects. J.Acoust.Soc.Amer., 31, 1606-1608.

THURLOW, W.R., and ELFNER, L.F. (1961) Localization effects with steady thermal noise in one ear and pulsed thermal noise in the other. Science, 134, 617.

TOBIAS, J.V. and SCHUBERT, E.D. (1959) Effective onset duration of auditory stimuli. J.Acoust.Soc. Amer., 31, 1595-1605.

TOBIAS, J.V. and ZERLIN, S. (1959) Lateralization threshold as a function of stimulus duration. J.Acoust.Soc.Amer., 31, 1591-1594.

TORGERSON, W.S. (1958) <u>Theory and methods of scaling</u>. John Wiley and Sons, New York.

TREISMAN, A.M. (1961) Attention and Speech. Ph.D. Thesis, University of Oxford.

TREISMAN, M. (1959) Computer Program A1002 - Probit Analysis. University of London Computer Unit.

TREISMAN, M. (1963) Auditory Unmasking. J.Acoust. Soc.Amer., <u>35</u>, 1256-1263.

VOELCKER, H.B. (1960) A decision theory approach to sound lateralization. In Cherry (1960), Information Theory.

VOELCKER, H.B.Jr. (1961) Studies of binaural perception of speech and noise. Ph.D. Thesis, University of London.

WALLACH, H. (1940) The role of head movements, and vestibular and visual cues in sound localization. J.Expt. Psychol., <u>27</u>, 339-368.

WALLACH, H., NEWMAN, E.B. and ROSENZWEIG, M.R. (1949) The precedence effect in sound localization. Amer.J.Psychol., <u>62</u>, 315-336.

WARD, W.D. (1961) Studies on the aural reflex: I. Contralateral remote masking as an indicator of reflex activity. J.Acoust.Soc.Amer., <u>33</u>, 1034-1045.

WETHERILL, G.B. (1963a) Sequential estimation of quantal response curves. J.Roy.Statist.Soc. Series B, <u>25</u>, 1-48.

WETHERILL, G.B. (1963b) Personal communication.

WHITWORTH, R.H. and JEFFRESS, L.A. (1961) Time vs. intensity in the localization of tones. J.Acoust.Soc.Amer., <u>33</u>, 925-929.

WIENER, N. (1948) <u>Cybernetics</u>. John Wiley and Sons, New York.

WILEY, R.L. (1963) Binaural fusion of artificially generated vowel-type sounds. Special Project Report, Dept. of Elect.Eng., Imperial College of Science and Technology, London.

WIND, M. (1956) Handbook of Electronic Measurements. Vol.2. Polytechnic Press, Interscience Publishers, New York.

WITKIN, H.A., WAPNER, S. and LEVENTHAL, T. (1952) Sound localization with conflicting visual and auditory cues. J.exp.Psychol., 43, 58-67.

WRIGHT, H.N. (1960) Measurement of perstimulatory auditory adaptation. J.Acoust.Soc.Amer., 32, 1558-1567.

YATES, F. GOWER, J.C. and SIMPSON, H.R. (1963) A specialised autocode for the analysis of replicated experiments. British Computer Journal 5, 313-319.

ZWISLOCKI, J. (1953) Acoustic attenuation between the ears. J.Acoust.Soc.Amer., 25, 752-59.

ZWISLOCKI, J. and FELDMAN, R.S. (1956) Just noticeable differences in dichotic phase. J.Acoust.Soc.Amer., 28, 860-861.

Addendum

LEVITT, H. (1964). Notes on Equipment. Telecommunications Laboratory, Imperial College of Science and Technology, London, S.W.7.

ZELUFF, V., and MARKUS, J. (1949). Ed. Electronics Manual for Radio Engineers. McGraw Hill, New York.