PERCEPTIONS ARISING FROM SIMULTANEOUS SIGNALS THROUGH TWO SENSORY CHANNELS

Floyd Edward Toole

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,
1965

ABSTRACT

In studies of perceptions arising from signals presented simultaneously through two sensory channels it was found useful to employ an experimental technique in which subjects assessed and reported the apparent extent of lateral displacements of aural and visual images which were perceived to arise from closely specified aural and visual signals. Curves of image positional judgements versus the relevant aural or visual signal parameters were used to provide a measure of the interaction of the signals in multiple sensory channels.

Signals simultaneously presented in aural and visual sensory channels were demonstrated to give rise to a number of cross-modal effects. It was shown that elucidation of these effects required further knowledge of the processes underlying judgements of acoustic images arising from the simultaneous combination of acoustic signals from the two (left and right) aural channels, and a suitable study was undertaken.

A method based on acoustic image positional judgements was found to be valuable in clarifying the nature of perceptions in complex image situations. With binaural repetitive acoustic transient signals, a number of coexisting binaural acoustic images could be independently identified and lateralized. Certain of these images were identified as being of tonal character and were clearly related to the lower harmonics of the repetitive transient signal. Other images, of impulsive character, were found to be closely associated with detailed features of cochlear dynamics.

A model of the basilar membrane was employed to enable observations of inferred time course of basilar membrane displacements in response to experimental

signals. Based on current knowledge of cochlear neurophysiology, an hypothesis about the mechanism of binaural interaction was developed and applied with some success to the prediction of results of specially designed psychoacoustic experiments and to account for results of significant earlier experiments of other workers.

CONTENTS

Page

APPENDICES

LOCATIONS OF DIAGRAMS AND TABLES

Figures

TABLES

GLOSSARY OF TERMS AND ABBREVIATIONS

Acoustic Image: the subjective impression of a spatially concentrated source of sound which is perceived to arise from the presentation of acoustic signals to both ears. Also known as binaural sound image, or simply, image. A related phenomenon wherein a sound is perceived to arise at the ear at which a monaural signal is applied is here always referred to as a monaural acoustic image.

Multiple images: simultaneously perceived acoustic images which have characteristic pitches or timbres which enable them to be independently identified.

Split images: simultaneously perceived acoustic images of identical or closely similar character which are perceived to occupy different spatial positions.

Residue Image: a term invented to describe acoustic images which were not perceived to alter position with interaural time delay and which were perceived to coexist with other binaurally created acoustic images which did. These images were commonly perceived in experiments with asymmetrically filtered or masked acoustic transients.

Conceptual Space: the subjective region in which binaural acoustic images are created and move.

Lateralization: the process of judging the extent of apparent lateral displacement of acoustic images which are intracranially perceived, as in binaural earphone listening.

Localization: the process of judging the extent of apparent lateral displacement of acoustic images which are perceived as arising externally.

Compromise Judgement: a judgement of apparent acoustic image position which by virtue of its relation to overall judgement groupings at a given value of ITD could be interpreted as indicating a balance of attention between two spatially distinct acoustic images (i.e. split images).

Sensation Level (SL): the intensity of an acoustic signal in decibels above its measured absolute threshold level.

ITD: interaural time delay.

IAD: interaural amplitude difference.

S_1, S_2: in experiments employing binaural single (A) versus double (B and C) repetitive transients, the acoustic images which were hypothesized to arise from binaural fusions of A and B (S_1) and A and C (S_2) transients.

A-B, C-D, etc. : in experiments employing binaural repetitive pairs of transients (A and D in the left-ear channel, and B and C in the right-ear channel), the images which were hypothesized to arise from binaural fusion of A and B transients (A-B), C and D transients (C-D), etc.

Symbols used in graphical plots of judged image position versus ITD: unless otherwise stated in the figure, a plotted open circle (o) relates to judgements of the dominant image, or to the dominant element of a split-image (from Chapter 5 onwards). In the same figure, a plotted cross (x) relates to judgements of the secondary element of a split-image. In figures where a number of experimental results are shown in superimposed fashion the preceding designations do not apply, and the significance of the various symbols is indicated.

ACKNOWLEDGEMENTS

The author wishes to express his deep gratitude to his supervisor, Dr. B.McA.Sayers, for his guidance, encouragement and advice throughout the course of the research work described here.

He is indebted to those of his colleagues who spent many hours acting as subjects in experiments. He is especially indebted to his wife who assisted in a large proportion of the experiments, as well as in the preparation and typing of this thesis.

The author also wishes to thank the Beaverbrook Foundation, and the National Research Council of Canada for financial assistance during this study.

CHAPTER 1
INTRODUCTION

The fields of communications, sensory psychology and sensory physiology are each closely involved with the matter of how information fed through multiple sensory channels combines to give rise to perceptions of various types. Of particular importance in many ways are perceptions which arise from signals presented simultaneously to two aural sensory channels or alternatively to the aural and visual sensory channels. In the latter case, there are many circumstances in which these two sensory channels operate in parallel fashion, conveying information about a single environmental event, as for example, in visual monitoring of lip movements while listening to speech. These two quite different types of information appear to be combined in such a way that the overall perceptual experience is not that of two separate phenomena.

In such cases the intelligibility of verbal communication may be demonstrably altered by the removal of one or the other of the signals (e.g., Sumby and Pollack, 1954); however, these effects are difficult to quantify in a way which can be meaningfully related to the actual signals involved.

More useful in this respect is the aspect of sensory interaction which, in this example, enables aural and visual information from a common source to combine in such a way as to establish a unified impression of direction. On the level of the individual senses, impressions of direction may be described in terms of the apparent spatial positions of perceived aural and visual images. These images, or apparent spatial sources of localized visual or aural sensation, appear to arise from another form of sensory combination -

that which is effected on information channelled through the two auditory channels (left and right ears) or the two visual channels (left and right eyes). The perception of aural and visual images, therefore, provides a convenient means of studying the laws of combination of information supplied through the two auditory channels and the two visual channels separately or at one time.

An illustrative experimental method which is useful in the study of sensory interactions of this type is that which relies upon an estimate made by a subject of the apparent spatial position of an image. Subjects presented with aural or visual signals which are perceived to give rise to images may be required to judge on some scale (say, azimuth angle) the apparent positions of acoustic and visual images as they appear separately or in combination in the subjective acoustic and visual 'spaces.' By suitable choice of aural and visual signals it is possible to achieve a quantitative measure of the relation between the positional assessments of the subjective images and the signal parameters which are important in altering the apparent positions of aural and visual images. With such information the simultaneously presented aural and visual signals may be independently manipulated such that they would give rise to atypical combinations of potential aural and visual image positions. Under these conditions it is possible to quantitatively assess the effect on the combined positional judgement of the contributory aural and visual signals, and thereby achieve a measure, in one respect, of sensory interaction.

A few workers have investigated sensory interactions using methods broadly similar to that just described. A number of interesting interactive phenomena were observed, viz., that where visual and aural cues were such that they indicated potentially different spatial sources, the visual cue dominated (see Jackson, 1953 and references cited therein). More interesting was the finding of Held (1955), that subjects who were exposed to uniformly non-corresponding aural and visual cues of localization showed a distinct tendency to adapt to the displaced 'acoustic environment'. The adaptation, which was such as to correct for the disparity in positional indications of the aural and visual signals, was demonstrated in purely aural localization experiments where some listeners indicated, for a single sound source, directions which were apparently the result of adaptation as well as directions which would normally be anticipated. The results were interpreted as having a bearing on the genesis of auditory localization.

However interesting these findings are it is difficult, because of the experimental techniques used, to relate the observations to quantitative measures of acoustic and visual signals. The purposes of the present study were, therefore, to conduct experiments based broadly on those referred to above, but in which the signals used were specified with precision sufficient to enable the psychophysical measurements to be related to the relevant physiological functions.

It is intended to report initially, certain preliminary experiments of a bimodal (cross-modal) kind which have been devised on the basis of simple experimental techniques. It will be argued that these experiments clearly indicate significant combinatorial effects, but particularly draw attention to the need for a more clear-cut understanding of basic combination of binaural auditory stimuli. This finding reinforces the significance of undertaking a study more closely associated than hitherto with the known facts of neurophysiology of binaural combination. Since all of these effects are mediated by neural mechanisms it seems now important to investigate if engineering techniques can be used to clarify neurophysiological mechanisms; any further success in this endeavour would clearly provide a basis for studying other forms of sensory combination, perhaps in the cross-modal case, using experiments of the type described below.

CHAPTER 2

THE EXPERIMENTAL TECHNIQUE

2.1.0 Introduction

The purposes of this study required, in effect, that three experimental methods be developed, one each for the aural and visual experiments, and one for the bimodal (combined aural and visual) experiments. However, it was expected that suitable design of the aural and visual experimental techniques would enable their simple combination in bimodal experiments. It was therefore important that these two methods be compatible in operational aspects as well as from the subject's point of view.

In brief, experimental methods were required whereby a subject's assessments of direction or position of acoustic or visual images could be quantitatively measured with equal facility whether these phenomena were studied in isolation or in combination.

The initial problems were concerned with the following general matters:

(1) <u>the technique of signal presentation</u>. A choice between 'dynamic' or 'static' signal presentation was required. In the former, the image positions would be altered while the subject monitored the signals continuously; whereas, in the latter case, the subject would not witness parameter changes. Although it might be of ultimate interest to know exactly how the various methods of presentation influenced localization performance, it was judged that the purposes of this study would be best served by using the static presentation technique. This was selected for several reasons:

(i) instrumentation and conduct of the experiments was greatly simplified because, in contrast to the dynamic method, strict control of the signal parameters at all times was not essential. The parameters were adjusted for each signal presentation and then ignored until the following presentation. In the meantime, the experimenter was free to record results, etc.

(ii) subjects had less immediate information of the history of image positions and therefore this factor might be of rather less importance in subsequent positional judgements.

(iii) the subject's task was eased by the fact that signals were presented only when responses were required. Further, response times were not important, and subjects could verbally report, position a pointer, etc., in the knowledge that the result would not be recorded until the task was completed.

(2) <u>the nature of the aural and visual signals</u>. Definable aural and visual signals were required which could be easily and independently controlled and which would be unambiguously and generally perceived as acoustic and visual images of unique spatial position for each value of the relevant signal parameter. Moreover, the aural and visual signals should be acceptable to subjects as conveying image positional information of comparable significance.

Close specification of acoustic signals applied to the two ears is most readily achieved with binaural earphones. The advantages of this method of acoustic signal presentation in conventional psychoacoustical experiments are well known; however, in the present experiments two difficulties arose. Firstly, in binaural earphone listening, acoustic images are most commonly perceived as being within the head. This is not only an unnatural situation, but also it occasions the second major difficulty, which is due to the fact that quantification of the apparent position of an intracranial acoustic image is not necessarily a natural talent. Other studies have largely avoided this problem by requiring listeners simply to estimate whether the acoustic image is left or right of the medial plane, or to have listeners adjust a signal parameter to cause the image to be moved to the apparent centre position. A third method, occasionally used, has been to have listeners adjust the position of a 'pointer' image to match that of a 'target' image. None of these methods was of any practical assistance here. It was therefore necessary to develop a scheme by which subjects could learn to judge and report with some precision the apparent position of an intracranially-perceived acoustic image. Such a procedure has been called 'lateralization'. It was then necessary to investigate if subjects could be trained to accept this positional information as being comparable and associated with the positional information from a visual display.

The general properties of binaural acoustic images are well documented, e.g., see von Békésy (1960), Licklider (1951), Woodworth and Schlosberg (1955). In brief, when signals applied to the two ears are identical in every respect, the resulting sound image (in earphone listening) is commonly reported as being near the centre of the head. If the left-ear signal, say, was caused to arrive later than the right-ear signal, there would exist an interaural time delay (ITD) and the acoustic image would be commonly perceived to move towards the ear receiving the earlier signal, the right ear. Similarly, if the amplitude of the left-ear signal, say, was increased relative to that of the right-ear signal, an interaural amplitude difference (IAD) would exist and the acoustic image would commonly be perceived to move towards the ear receiving the louder signal, the left ear. These basic rules hold for most common signals; however, there are signals for which one or the other of these parameters (ITD or IAD) is more effective in altering the apparent position of the acoustic image.

The choice of acoustic transients (commonly referred to in the literature as 'clicks') as a signal was based largely on the observation that the acoustic image which was perceived to arise on the binaural presentation of such signals was apparently unambiguous and was perceived to move in what seemed to be a well-defined and unique manner under the influence of ITD. As ITD was increased from zero the image was perceived to move monotonically to one side, reaching an apparent limiting position near the ear at an ITD of about 0.6-1.0 msec and remaining there as ITD was increased further. It seemed that ITD variations of ± 0.6-1.0 msec would be suitable to enable this binaural signal to fulfill the requirements of this study. The brief electrical pulses which were applied to the earphones were convenient to generate and ITD variations relatively easy to instrument.

The visual signal presented fewer problems as it was judged that a vertical line or spot displaced horizontally on an oscilloscope screen would suffice. Such a signal was easily generated with standard laboratory equipment.

(3) <u>other factors</u>. Decisions concerning the method by which subjects reported positional judgements, the mode of experimental conduct and the means of experimental data acquisition were all closely coupled. It was necessary that the subject's quantitative responses be in a form which could be rapidly interpreted and recorded. In manually conducted experiments the responses

could be verbal and in numerical form. However, many experiments were conducted under instrumented programme control in fully automatic fashion and here it was necessary that the responses be in a form which could be encoded, sampled and recorded using standard data processing apparatus. In order that the response indication be compatible with either manually or automatically conducted experiments the response indicator finally adopted was in the form of a linear graduated scale against which subjects compared their assessments of image displacement. Subjects either verbally reported the appropriate numerical scale value (manual) or aligned a pointer with the appropriate graduation (automatic). In the latter case, the pointer was mechanically coupled to a potentiometer shaft, the output voltage of which (linearly related to the scale calibrations) could be readily measured and encoded using a digital voltmeter, or otherwise utilized.

In the case of visual image localizations it was clearly not desirable that subjects should be able to readily compare visual image position with response scale calibrations. To avoid this, the visual signal display and response indicator scale were positioned such that a transfer of attention from one to the other involved a shift of focus and angle of view.

The fully automated and programme-controlled experiments mentioned above were used in developmental stages where experiments were aimed at evaluating the general effects and utility of various gross or subtle changes in experimental technique. As well as assuring the strict reproduction of experimental signal presentations on as many occasions as required, the method left the experimenter free to observe the course of the subject's responses as they were reported. The latter was often instructive.

2.2.0 Apparatus

Automation of these experiments necessitated the construction of a number of pieces of specialized apparatus. The heart of such a system is the programmed central control, in this case a four-channel magnetic tape unit (see Appendix A2). Recorded on the programme tapes were signals which would independently control the auditory and visual signal parameters, the data print-out apparatus, and the presentation and removal of auditory and visual signals. These signals, which were recorded using a pulse-width modulation technique, were in the form of D.C. voltages of values corresponding in a linear fashion to the desired image position (visual) or the desired ITD value (aural). Control of the ITD parameter with such voltages required special equipment for generation of the binaural pulse signals (see Appendix A1).

The experimental data print-out involved the simultaneous automatic plotting of a positional judgement versus ITD graph and the punching of a paper tape containing coded numerical indications of positional judgements. These punched-tapes were subsequently fed into a teleprinter which produced a typewritten numerical record of the experimental results. They also had the potential of being used as data tapes for computer assessment of results. For details of this apparatus see Appendix A3.

The overall system for the conduct of automated experiments is illustrated in Fig. 2.1. Voltage outputs 1 and 2 from the magnetic tape unit were used to control ITD (aural) and X-axis displacement (visual) respectively. Output 3 carried a simple two-level voltage which was used to control the presentation of the aural and visual signals. The binaural pulse signals were switched on and off with voltage—controlled transistor gates having a gradual turn-on (about 50 msec rise time). The visual signal was brought into view by the application of a steady potential to the Y-axis input of the display oscilloscope. In the absence of this voltage the image was displaced off the screen and out of view. During the presentation, which could involve either aural and visual signals, or both, the subject would adjust the response indicator to the appropriate value. At the end of the presentation the removal of the Output 3 voltage was sensed by the "Sample Control" and a 'read' signal was applied to the digital voltmeter. The latter digitized the voltage output of the response indicator and initiated the encoding and punching operations. The "Sample Control" also applied a signal to the X-Y plotting device which caused the pen to drop onto the paper for a moment and plot a point on the graph reflecting the X and Y voltage conditions, i.e., visual image position or ITD (dotted alternative connection) and positional judgement, which prevailed at that time. Between successive signal presentations the voltages at outputs 1 and 2 were altered, thus establishing new parameter values for the next judgement. The horizontal and vertical dimensions of the visual image on the oscilloscope screen were established by noise generators applied to the X and Y oscilloscope inputs respectively.

Later experiments were manually conducted and used the modified scheme as illustrated in Fig. 2.2. Here, the repetitive pulse signals were controlled by a clock generator which established the repetition period (10 msec) and which after a fixed delay time (Delay 1) initiated the signal pulses which were applied to the right earphone. A non-delayed output triggered a monostable multivibrator which was used to gate a sweep (linear voltage ramp) generator, the output of which was applied to a voltage comparator. When the voltage of the ramp equalled that of the second input to the comparator (from stepped potentiometer 1), a trigger pulse was produced which initiated the signal pulses which were applied to the left earphone. Thus, by altering the voltage applied to the comparator from the stepped potentiometer the temporal relationship between the left- and right-ear pulses (i.e., ITD) could be altered. The same system in a different form was used in the automated experiments and is described in more detail in Appendix A2.

Coupled to stepped potentiometer 1 was a second similar potentiometer. The variable voltage output of this was used as an X-deflection signal on the display oscilloscope (visual image positioning). As in the automated experiments the nature of the visual image was established by horizontal and vertical noise deflections.

Both aural and visual signals were presented and removed as before, only here the control was manual and separate for each signal.

In both automated and manual experiments there was provision for introducing a fixed ITD bias. This was effected by altering the fixed delays which, in each case, were established in the right-ear channel, and which normally were set equal to the variable delay in the left-ear channel when this was adjusted to the desired centre position.

2.3.0 Experimental Technique

2.3.1 Experimental Technique: Auditory Image Lateralization

As was stated above, the fundamental problem was that of training subjects in the task of judging the apparent lateral position of an intracranially-perceived acoustic image. Jeffress and Taylor (1960) have shown that listeners can perform this task when the positional indication is that of azimuth angle. Von Békésy (1960) and Guttman (1962) have mapped the course of apparent acoustic image position with ITD but in a fashion which was not precisely quantitative. In short, there was very little background information of practical use. The early experiments were, therefore, of an exploratory nature generally aimed to test the practicality of a number of schemes and to evaluate the general requirements in the way of auditory signals.

Listeners confidently reported simple 'left' or 'right' judgements, indicating that the perceived acoustic image was to the left or right of what was judged to be the central position; however, it was difficult for them to find a basis on which to judge the amount of the apparent lateral displacement as ITD was introduced. Most listeners, however, were able to 'follow' the apparent lateral position of the sound image by positioning a pointer or lever on a real or imaginary scale which in a simple way reflected the acoustic 'space'. This was spontaneously done and reported by a few listeners who said that they projected onto this scale not only the lateral physical extent of the head (i.e., separation of the ears and apparent limit of image displacement) but also the nose (sic), eyes(sic) etc., to which they said they compared the apparent position of the acoustic image. In other words, they had used this external scale to parallel their impressions of acoustic image displacements.

The acoustic signals used were those which resulted from the application of brief electrical pulses (duration 0.4 msec) at a repetition period (T) of 10 msec to each of a pair of S.G. Brown Super K earphones. This transient signal was chosen as one of many which were judged by listeners as giving rise to apparently clear-cut acoustic images. Signals were presented at what was judged a comfortable listening level, measured as being about 50 dB SL. Sound levels were adjusted and checked at regular intervals using a calibrated artificial ear (Bruel and Kjaer, type 4109 with 9 cc coupler). It was found that over the range of signal levels likely to be used in forthcoming experiments (20 dB SL to 70 dB SL), there was a closely linear relationship between the amplitude of electrical pulse inputs to the earphones and the measured acoustic outputs. Therefore, for convenience, the signal levels were referred to the amplitudes of the electrical pulse signals.

As discussed above the signals were presented in the 'static' fashion, i.e., with fixed ITD's which were altered only after the signals had been removed. The duration of signal presentations, of the intervals and of the overall experiments were determined by several factors, viz.

(1) the time required for the listener to make a positional judgement and to indicate the result,

(2) the number of such responses he could make before fatigue caused a degradation of performance,

(3) the number of fixed ITD values chosen for study,

(4) the number of judgements at each ITD value required to enable a realistic assessment of the sub-

ject's performance.

Previous experience had indicated that most subjects could complete their responses within a period of 4 seconds. It had also shown that if experiments were longer than about 15-20 minutes subjects became restless and unable to concentrate effectively on the task. It was decided therefore that signal presentations of 4 seconds duration, followed by 4 seconds of silence (during which parameter changes were made) would be convenient. An experiment of about 15 minutes duration allowed 110 judgements which were equally distributed among 11 ITD values (10 judgements each at ITD = 0 and 5 plus and 5 minus values). Including instructions to the subject and other preparations a typical experiment may have lasted 25 minutes.

The sequence in which signals with the 110 values of ITD were presented was determined by the random selection of ITD values from the pre-established set of 110. Three programme tapes were recorded having different randomized sequences and these were used interchangeably in the experiments so as to minimize the effects of learning.

Various devices were tried as positional judge ment indicators in these early experiments. The first really satisfactory method of response indication was that in which listeners adjusted, by means of an unadorned knob, the horizontal position of a spot on the oscilloscope screen so as to match their impressions of acoustic image position. This technique was useful for development purposes only, because ultimately the oscilloscope would be used for the visual signal display.

Prior to experiments, listeners were presented with acoustic signals having: ITD = 0, ITD = + maximum and ITD = - maximum, each repeated five times in systematic order. It was believed initially that this would provide listeners with a sense of the limits of image motion so that positional judgements on subsequent randomized signals could be appropriately scaled. Listeners were not told how many values of ITD were used, but it was impossible to keep this information from eventually reaching some of them.

Results typical of these early experiments are illustrated in Fig. 2.3. Here are shown for each category of ITD, individual and averaged positional judgements as indicated by two listeners. Fig. 2.3(a) illustrates a result obtained with an experienced listener (one who had participated in the preliminary experiments). Fig. 2.3(b) illustrates an example of results obtained with a naive listener. The example shown was the first trial of this subject. There is a clear difference in the linearity of the relationship between averaged positional judgements and ITD, which in these cases was altered over the range ± 1.0 msec. In fact, the degree of this non-linearity was found to vary considerably among listeners; it was noted, however, that the more experienced listeners tended to produce the more nearly linear averaged positional judgement versus ITD curves.

The results described above were obtained using a response indicator with only a centre marker. Subjects were free to adjust the position of the spot to any position within the allowed range. However, categorized responses would be more convenient in a number of ways, e.g., data handling would be eased because of the limited number of response values and, in manually conducted experiments, verbal responses could be simple numerical indications of categories. To test the utility of this scheme, the oscilloscope screen was fitted with a scale calibrated with a centre marker and five equally spaced markers on either side. Subjects were required to align the spot on the oscilloscope screen approximately with one of these graduations (exact alignment was not required because of the categorization nature of the response). Representative results obtained using this scheme, are illustrated in Fig. 2.4.

Fig. 2.4(a) illustrates in superimposed fashion, results of two similar experiments using the listener employed previously to obtain the result of Fig. 2.3(b). Fig. 2.4(b) illustrates two results obtained with a naive listener. The results were obviously quite satisfactory. The relation between averaged positional judgements and ITD was nearly linear in each case. Although no special significance is attached to this finding, such a relation is clearly more convenient to work with than, say, the curve of Fig. 2.3(b). Further, response indications were made more rapidly, thereby allowing more time for the important task of lateralizing the acoustic image.

As a consequence of those findings a new form of response indicator was introduced: a set of 11 push-buttons arranged in a straight line over a span of about 8 inches. The underlying switching mechanism was interlocked such that the pressing of a button cancelled the previous response and latched the new indication. Electrical contacts associated with each push-button, when closed, caused a voltage to appear at the output which was a function of the location of the depressed button.

Tests indicated that this response indicator retained the apparent advantages of its predecessor. Experimental results were usually similar to those of

Fig. 2.4; however, certain subjects continued to demonstrate non-linear response curves, although more linear than that of Fig. 2.3(b). The shape of the positional judgement versus ITD curves appeared not to alter seriously from day to day.

A further factor which was subject to some variability was the ITD at which the judged image trajectory was seen to cross the subjective centre axis. The limits of ITD variability found here were about $\pm$ 100 μsec. The largest differences were more often observed between listeners than between occasions with the same listener. Experienced listeners often demonstrated closely reproduced values of centered ITD's over a period of a week or more.

It was judged that at this stage the results of acoustic image lateralization experiments, while perhaps not ideal, were sufficiently reproducible to justify application of the present technique to the more complex bimodal experiments.

2.3.2 Experimental Technique: Visual Image Localization

The task of identifying the position of a visual image is clearly quite different from that of identifying the apparent position of an intracranially-perceived acoustic image. Listeners required some special preparation before they could judge,with any degree of reliability, the apparent degree of lateral displacement of an auditory image; however, in contrast, subjects were able to judge the lateral position of simple visual (vertical line or spot) image on an oscilloscope screen immediately and with small error. While in isolation this might have been interpreted as admirable, it meant that this simple and obvious technique was clearly unsatisfactory for application in bimodal experiments.

In other words, in bimodal experiments, the relative simplicity of the visual localization task would be expected to result in the almost complete domination of the visual information.

The problem was, therefore, to alter the visual presentation so as to make the task of positional judgement more difficult. Two important factors which assisted judgements were: the clarity of the image itself, and the availability of various landmarks for use by the subject as points of reference, e.g., the bounds of the oscilloscope screen, and knobs, switches, etc. in the vicinity.

The scheme by which visual landmarks were minimized was simple in concept, but required the construction of an experimental booth and special fixtures. The large screen display oscilloscope was placed behind an unmarked mask having wide (4 ft.) horizontal slit through which the subject viewed the screen. Suitable light baffles and an inclined sheet of tinted perspex prevented reflections (possible reference points) and darkened the slit so that all that could be seen was the luminous image. See Fig. 2.5. The assembly was placed within an acoustically treated booth, which was provided with a constant low-level of ambient illumination. Heat generated by the apparatus was removed by forced air ventilation. The booth provided a uniform and apparently innocuous environment for the conduct of experiments.

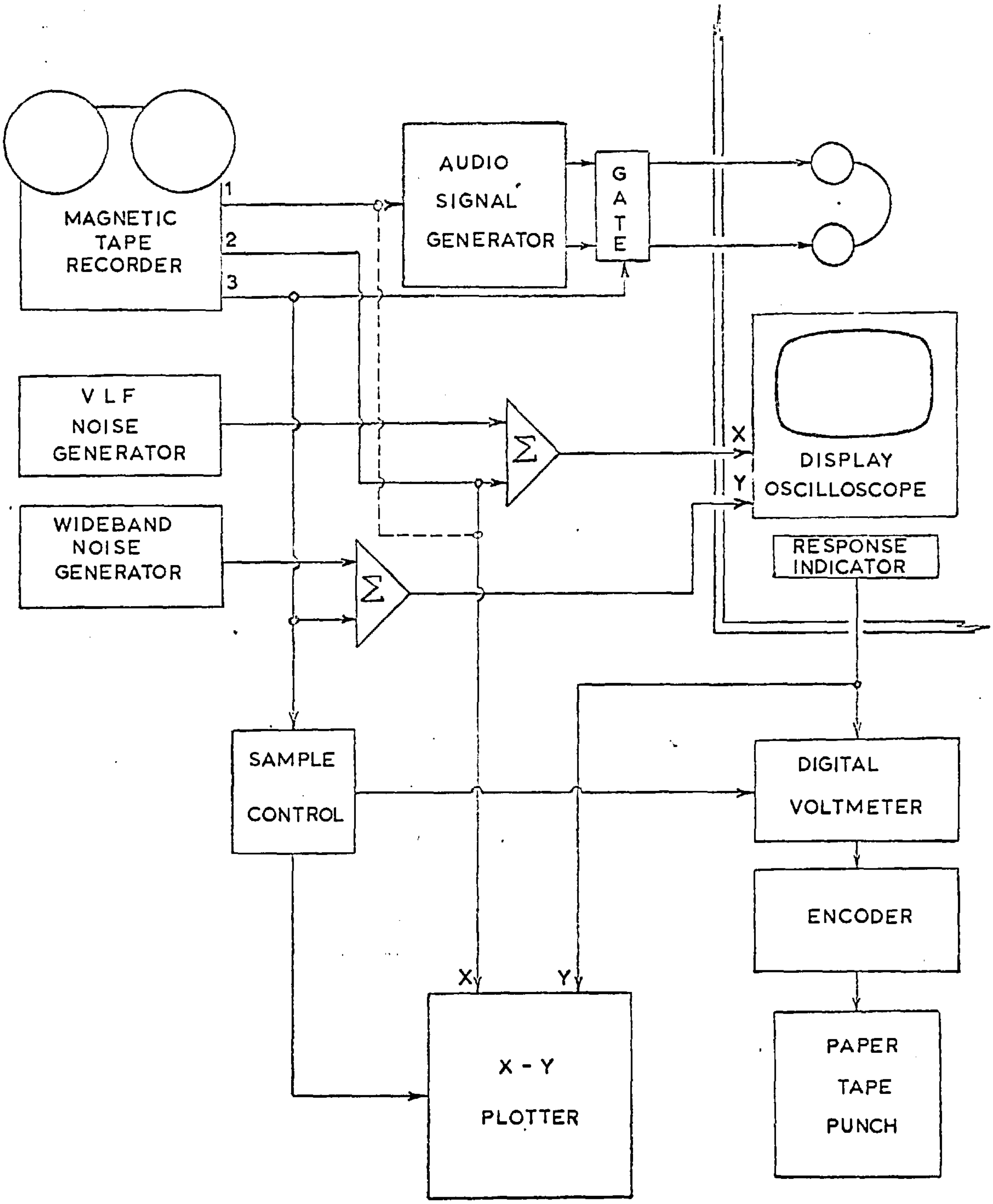
The character of the visual image (a 4 cm long vertical line) was altered by causing it to be rapidly displaced from side to side in a random fashion. The amount of peak lateral deflection and the maximum rate of movement were established by trial and error. The maximum rate of movement was governed by the upper spectral limit of the low-frequency random noise signal which was applied to the X-axis. It was found that if the upper frequency limit was too high, the persistence of the phosphor on the oscilloscope screen produced a steady image with a clear centre; however, if the upper frequency limit was too low, the image moved so slowly that it was impossible to ascertain a central tendency of the random image motion within the presentation time (4 seconds). A reasonable compromise was judged to be an upper spectral limit of about 15 cps.

The amplitude of random horizontal motion (i.e., the horizontal dimension of the image) was largely governed by the horizontal span of static image displacements (i.e., image positions). There seemed to be no uniquely satisfactory value for the latter, and the convenient value of 10 cm was chosen; thus, the interval between adjacent image positions was 1 cm. The horizontal dimension of the image was established by observing the scatter of positional judgements for various image widths. The aim was to achieve judgement scatter comparable to that observed in results of acoustic image lateralization experiments. It was assumed that the scatter of positional judgement was in some sense a measure of the difficulty of the task; introspective comments of the subjects generally confirmed this assumption. The horizontal dimension finally chosen was about 4 cm.

With the equipment as just described, subjects were given no fixed references for positional judgement other than their bodily positions, and this could not be expected to be very stable (the image moved within a viewing angle of about 10°). Subjects' reports and experimental results verified this. This condition was improved by the provision of a small (1/16 inch diameter) illuminated reference dot centered about 6 inches above the horizontal plane of image motion, and in the same vertical plane as the image (to prevent parallax).

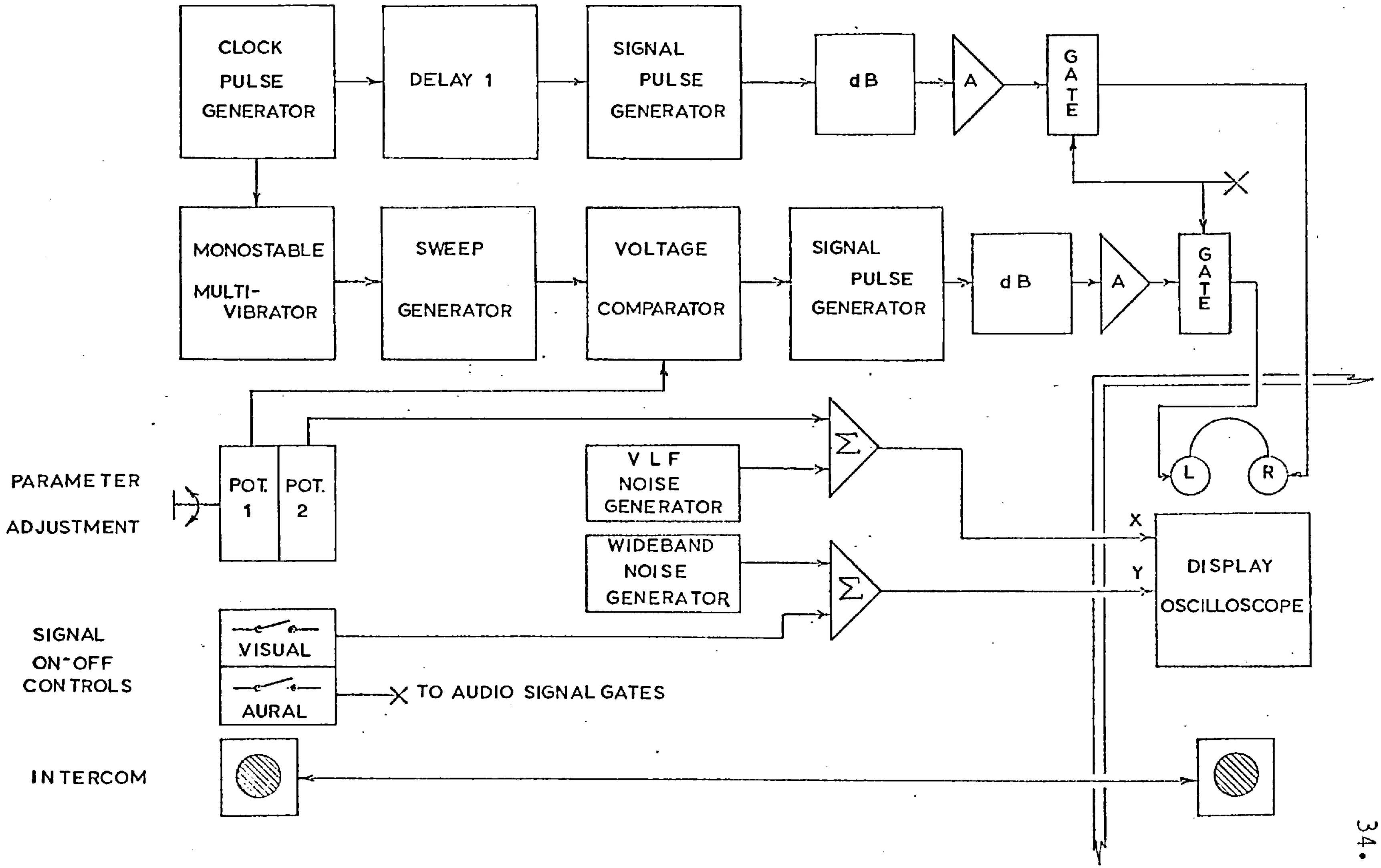
This served to stabilize the pattern of judgements without seriously affecting the apparent judgement precision.

Judged image trajectories which resulted from experiments conducted with the apparatus described above were generally linear, with judgement scatter only slightly less than that observed in acoustic image lateralization experiments.



Arrangement of Apparatus for Automated Experiments

FIGURE 2.1



Arrangement of Apparatus for Manually Controlled Experiments

<u>FIGURE 2.2</u>

FIGURE 2.3(a)

SUBJECT: F.E.T.

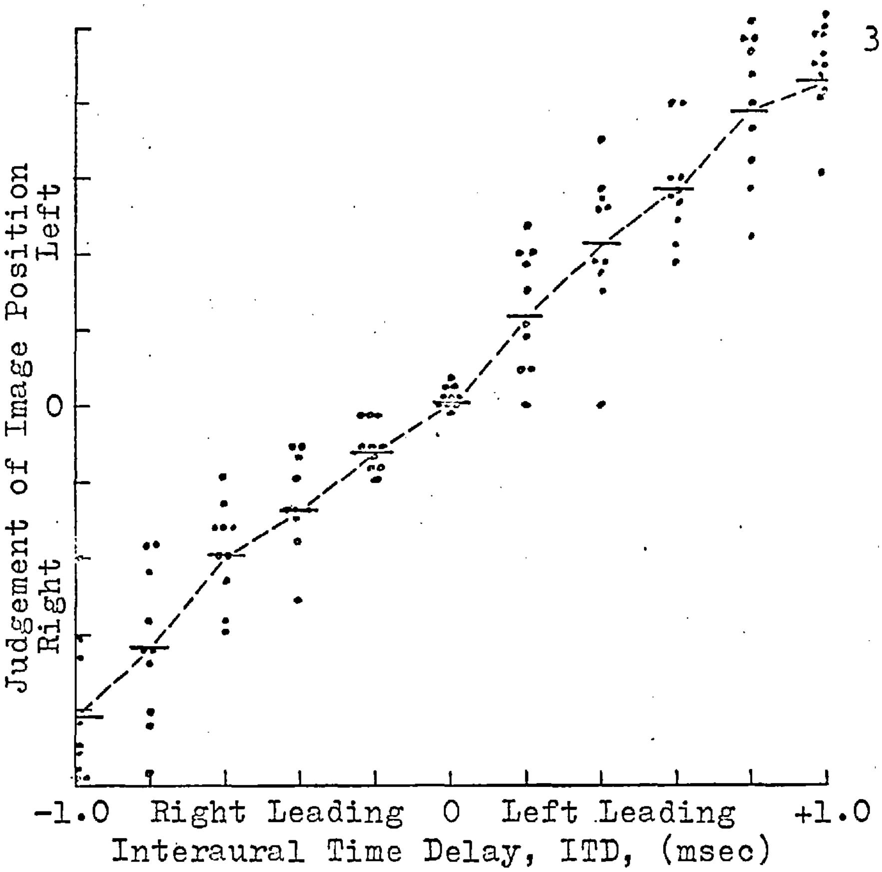


FIGURE 2.3(b)

SUBJECT: N.B.T.

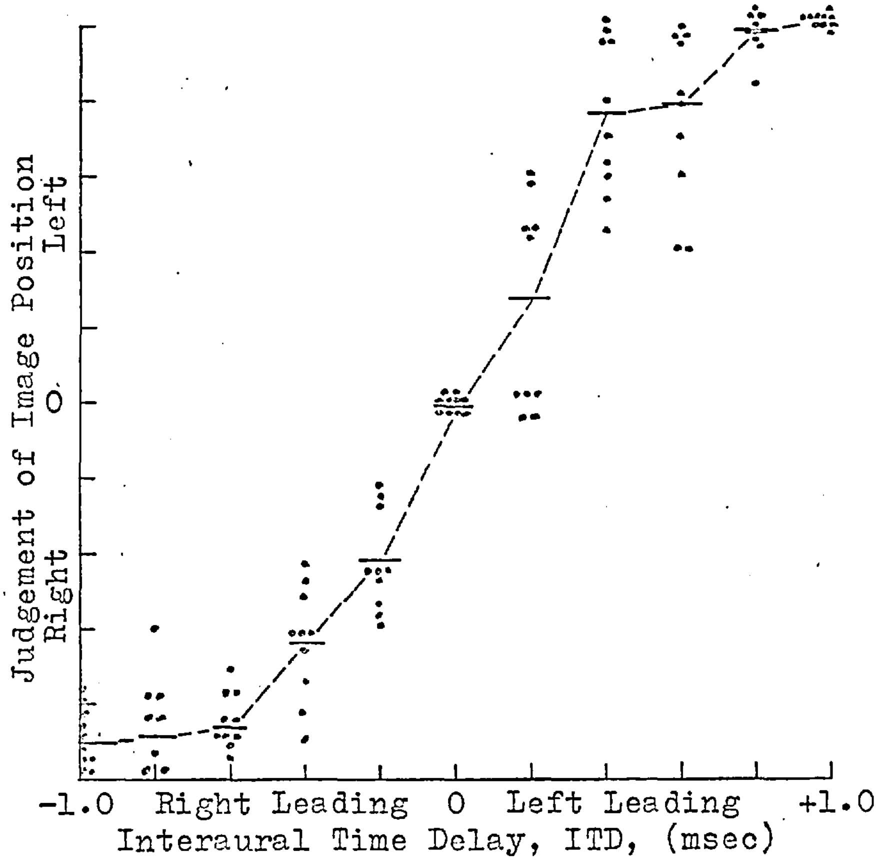


FIGURE 2.4(a)

SUBJECT: N.B.T.

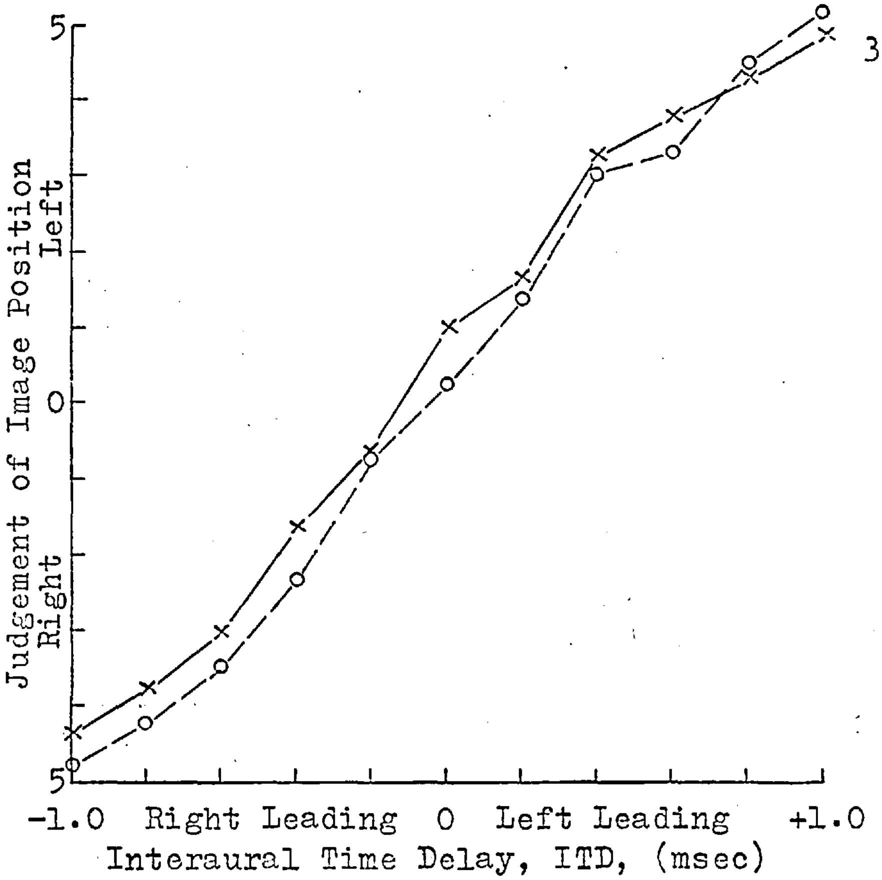
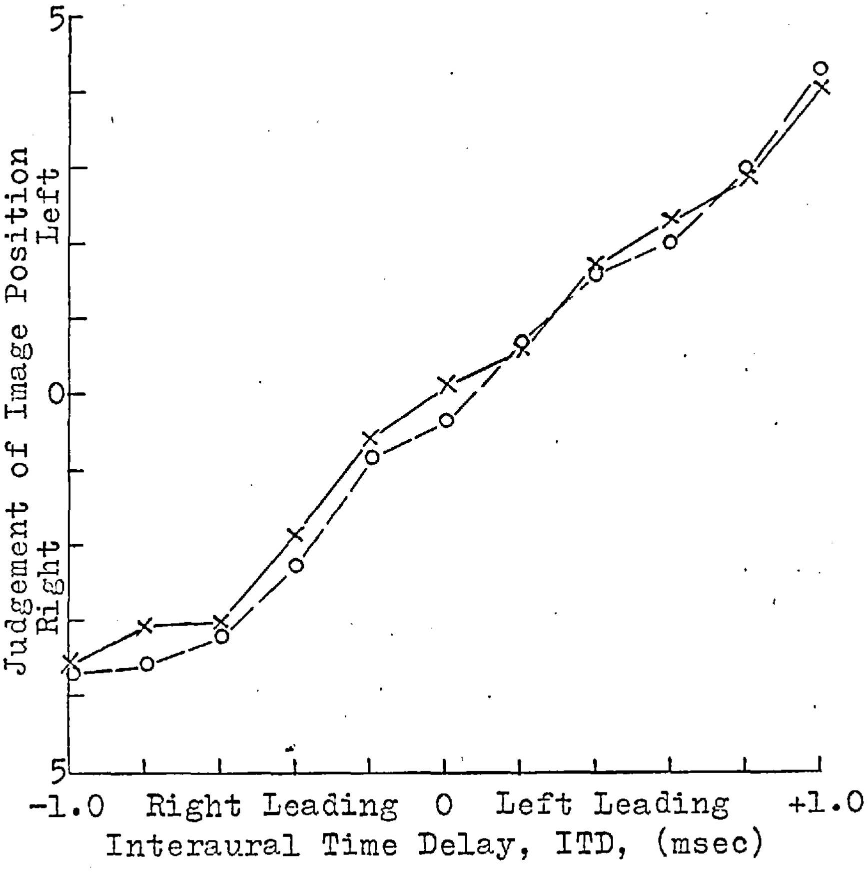


FIGURE 2.4(b)

SUBJECT: G.A.



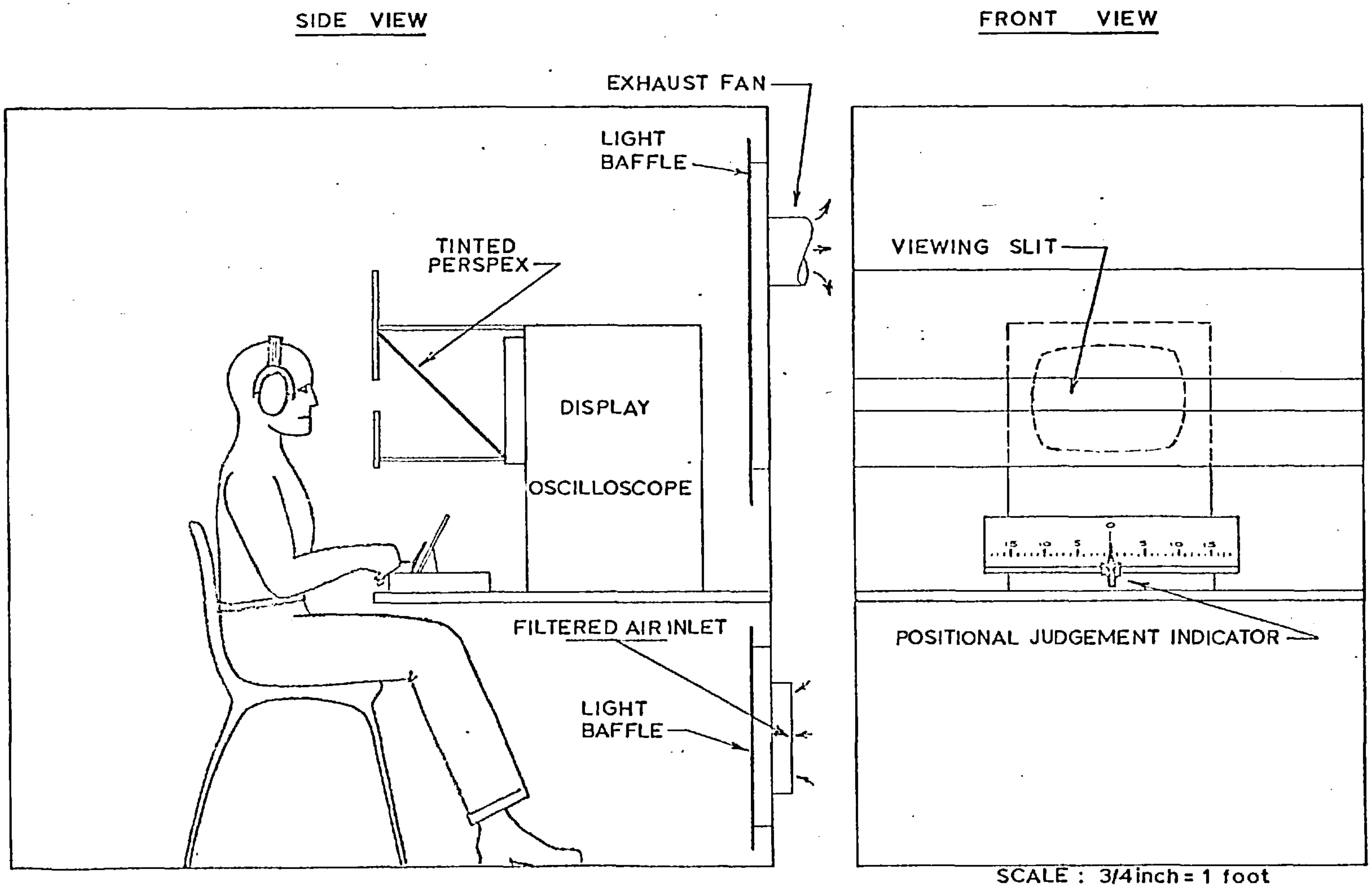

SIDE VIEW
FRONT VIEW
EXHAUST FAN
LIGHT BAFFLE
TINTED PERSPEX
DISPLAY
OSCILLOSCOPE
FILTERED AIR INLET
LIGHT BAFFLE
VIEWING SLIT
15 10 5 0 5 10 15
POSITIONAL JUDGEMENT INDICATOR
SCALE : 3/4inch = 1 foot

FIGURE 2.5

CHAPTER 3

A PRELIMINARY STUDY OF CROSS MODAL EFFECTS IN VISION AND AUDITION

3.1.0 Introduction

The methods of acoustic image lateralization and visual image localization as described in the previous chapter appeared to be sufficiently tractable to enable their application in the more complex bimodal experiments where single positional judgements were to be based on simultaneous presentations of aural and visual signals. Background experiments using these two methods had shown to give, for a number of subjects, results which were closely reproducible, although not always closely similar. However, the differences (in the linearity of judgement curves) would be expected to raise difficulties primarily in comparisons of results with different subjects.

It was first necessary to ascertain whether or not cross-modal effects could be elicited and, if so, to identify and seek a measure of such effects. An important aspect of such a study would be to ascertain the relative importance of the aural and visual information, to the combined positional judgements. To this end, sequences of experiments were arranged in which the lateralization potentials of the aural and visual signals respectively were caused to differ. It was postulated that where the positional indications of the two signals were identical there would be no important modification to judgements as would pertain to either signal alone; however, where a positional disparity was introduced, it was thought that the judgements may pertain either to the aural or to the visual signal depending on which was in some sense the more important, or perhaps to some intermediate position indicating a compromise.

3.2.0 Bimodal Experiments: Series 1

It was considered to be of fundamental importance that subjects should consider the aural and visual positional information as being in some sense associable. As a background to the bimodal experiments, therefore, subjects were practiced in the tasks until they were thoroughly accustomed to reporting, by means of the push-button response indicator, the apparent lateral displacements of aural and visual images. All of the experiments were conducted under programme control with identical preparatory periods during which the images were repeatedly presented in their left and right maximum and (nominal) centre positions; see section 2.3.1 and Appendix A2.

With these results as standards for comparison, each of the listeners was put through a prearranged sequence of unimodal and bimodal experiments.

The first set of these experiments proceeded in the following sequence:

(1) bimodal: aural and visual signals normal, i.e., which individually give rise to perceived maxima of image displacements closely symmetrical about a judged central position (normally associated with ITD = closely 0 or a physically centered visual image). Aural and visual images would be expected to be perceived as being displaced by proportionally similar amounts from the medial position.

(2) bimodal: visual signal normal, aural signal presented with a fixed ITD bias such that, on the basis of unimodal image lateralization results, the acoustic image would be expected to always have an apparent position to the left of that normally judged for the visual image. The amount of the positional disparity was estimated to be about 1/10 of the total span of

image displacements, i.e., a fixed ITD bias of 0.13 msec (left leading).

(3) bimodal: visual signal normal, aural signal presented with a fixed ITD bias as in (2) above. The amount of the bias was increased to 0.26 msec (left leading).

(4) acoustic image lateralization: aural signal as in (3) above.

(5) acoustic image lateralization: aural signal normal.

The experiments were conducted in succession over a period of four hours. The subject (female) was given instructions as for the regular aural or visual experiments, but in addition was told that on these occasions the signals would be presented simultaneously. No hint was given as to the real purpose of these experiments, nor of the signal manipulations. In fact, the subject thought that these were only preliminary experiments and commented that she had considered the signals as conveying equivalent positional information.

Fig. 3.1(a) illustrates, for comparison, results of two acoustic image lateralization experiments and one visual image localization experiment conducted earlier with the same subject. Fig. 3.1(b) illustrates in superimposed fashion the results of experiments described in (1) to (4) above. The results have been plotted in this way to demonstrate the close similarity of these findings. In general, it seems that the positional judgements in the bimodal experiments were dominated by the visual signal, no changes being apparent from the progressive modification of the aural signal. The subject's introspective comments indicated that she had not detected the disparity. Moreover the acoustic image lateralization result (E7) so closely resembled the bimodal result as to suggest that the sequence of bimodal experiments may have caused a change in the acoustic image lateralization behaviour of the subject (cf. E1 and E2). A test of this is provided in the plotted results of Fig. 3.2(a), which includes the acoustic image lateralization result (with fixed ITD bias = 0.26 msec) which was just described (E7) and the result of the normal acoustic image lateralization experiment (no ITD bias) which followed it (E8). The results have been plotted to indicate averaged positional judgements for corresponding ITD's in the two cases. It is clear that the positional judgements have been systematically altered over a large portion of the ITD range, and while this may not be entirely unexpected in regions away from ITD = 0, it is somewhat surprising that this should occur near ITD = 0. In other words, it would appear that the subject's subjective auditory centre has been effectively displaced. As a measure of the significance of the shift the standard statistical 't' test of significance of difference between two sample means was applied to results in the vicinity of subjective centre. Results of the test, which are included in Fig. 3.2(a), indicate that for the points examined, the differences are significant to at least the 0.05 level (i.e. differences equal to or greater than those observed might be expected to occur in less than 5 out of 100 random samples).

To check that this result was not fortuitous, the same experiments were repeated on the following day, with closely corresponding results. In fact, four weeks later results of an isolated acoustic image lateralization experiment using this listener were of similar form. The results of four acoustic image lateralization experiments in which the signals were normal and three in which the signals had a fixed ITD bias were collected and separately averaged. The averaged results of these experiments, which were conducted over a period of eight weeks, are shown in Fig. 3.2(b). The ranges within which the averaged results fell are indicated in the figure. It will be seen that the two averaged curves are of the same form as those of Fig. 3.2(a). In a sense, these curves confirm the remarks about the curves of Fig. 3.2(a); however, the possibility arises that bimodal experience may not be the sole factor underlying these apparently different performances. This may be deduced from the similarities of results of post- bimodal and pre-bimodal experiments with similar (normal) signals.

Consider next the results of a similar sequence of experiments conducted with a different subject (H.F., male)Fig. 3.3(a) illustrates comparative results of experiments using normal aural and visual signals. These results are of special interest because of their similarity to those of N.B.T. which are illustrated in Fig. 3.1(a). To assist in the comparison, the acoustic image lateralization results of these earlier experiments (E1 and E2) have been replotted here.

Fig. 3.3(b) illustrates in superimposed fashion, the results of the sequence of three bimodal experiments and the final acoustic image lateralization experiment as were described in (1) to (4) above. As in the previous case, Fig. 3.1(b), the results of the bimodal experiments (E14, E15 and E16) were all closely similar; however, the curve of the acoustic image lateralization experiment (E17) differed substantially.

Fig. 3.4(a) illustrates, for comparison, judged acoustic image trajectories for the post-bimodal experiment (E17) with the modified signal, and for the pre-bimodal experiment (E12) with the normal signal. In contrast with the comparable curves of Fig. 3.2(a) the differences between these curves are small and inconsistent.

For further comparison Fig. 3.4(b) illustrates the results of experiment E17 along with those of three comparable experiments conducted with N.B.T. The result of the present experiment (E17) is clearly quite different from the three basically similar results of the earlier sequence (E7, E10 and E11).

Observations based on these two sequences of experiments may be summarized as follows:

(1) in bimodal experiments positional judgements were clearly dominated by the visual signal information,

(2) in bimodal experiments where the aural signal was such that there would be a disparity in the potential positional information conveyed by the aural and visual signals, the disparity was apparently not detected by the subjects.

(3) one subject (N.B.T.) demonstrated an apparently modified form of positional judgement versus ITD relation after the bimodal experiments (Fig. 3.2a). However, it was not clear whether the effect, which could be described as a displacement of auditory centre, was attributable to the bimodal experience or simply to the asymmetrical (about ITD = 0) acoustic signal presentations. Whatever the cause, the effect was demonstrably present over a period of one month (Fig. 3.2b).

(4) the second subject (H.F.) produced rather different results. It was clear that the experimental manipulations did not seriously alter the ITD-to-positional-judgement correspondence near ITD = 0 which was demonstrated earlier by this subject (Fig. 3.4a)

This was confirmed by comparisons of results for this listener with those of N.B.T. (justified on the basis of the initially similar forms of response, Fig. 3.3a); see Fig. 3.4(b).

As a check on what might be anticipated from experiments in which the acoustic signal was modified in the absence of the visual signal, a further sequence of experiments was conducted. Following the basic pattern of the preceding sequences, a subject (female) participated in normal aural and visual experiments and in aural experiments where the signal had a fixed ITD bias.

Fig. 3.5(a) illustrates results of the normal acoustic image lateralization experiment (E18) and of two aural experiments in which a fixed ITD of 0.26 msec (left leading) was introduced (E19 and E20). The judgement curve of experiment E18 is almost perfectly linear, in contrast to those of the earlier subjects (experiments E1, E2 and E11 - see Figs. 3.1a and 3.3a). The principal effect of introducing the fixed ITD bias was to displace the judgement curve along the ITD axis such that ITD = 0 was still judged as producing a centered image. However, other effects may be seen and it is believed that these relate to characteristics of subjective scaling. On the one hand, judgements at the left-hand extreme positions appear to have been somewhat restricted and the linear extrapolation of the judgement curve which might have occurred has been prevented by the absence of push-buttons beyond that which indicates position 5 left. On the other hand, the judgement scale seems to have been expanded on the right-hand side apparently so as to utilize the full scale (introspective comments of the subject confirmed this observation).

At this point the subject participated in a single bimodal experiment (normal visual signal - aural signal with fixed ITD bias of 0.26 msec, left leading), followed by an acoustic image lateralization experiment with the same signal. Fig. 3.5(b) illustrates the result of the bimodal experiment (E21) along with that of a normal visual localization experiment (E22). The two results are closely similar, indicating the clear domination of the visual signal. For reference, a curve representing an average of the two aural experiments using the same signal (E19 and E20) is shown; this curve represents acoustic image positional judgements which would be anticipated in the absence of the visual signal.

The judged image trajectory for the final acoustic image lateralization experiment (E23) is illustrated in Fig. 3.6(a). Judgements have been clearly altered by the subject's experience in the bimodal experiment, but in a way which is not easily accounted for. For comparison the averaged results of aural experiments E19 and E20 (both with ITD bias) have been shown and also the result of the normal aural experiment E18. The latter two curves are strikingly similar in spite of the fact that one (E18) represents results with a normal aural signal, and the other (E19 and E20) represents results with an asymmetric (about ITD = 0) signal presentation. The characteristics of the result of experiment E23 are therefore even more significant and would appear to have resulted from some aspect of the subject's experience in the bimodal experiment. The direction and magnitude of the change are such as to escape simple interpretation.

3.2.1 Discussion of Series 1 Experiments

It is immediately clear that there is a form of sensory interaction which may be elicited in experiments of the simple type just described. However, certain of these effects are not amenable to immediate interpretation. In particular, it was generally observed that where aural and visual signals were simultaneously presented, positional judgements were dominated by the positional information of the visual signal. In no case did subjects report detection of any disparity between the positional information of aural and visual signals where such a disparity would be expected to exist (due to signal modification and comparisons based on unimodal image lateralization and localization results).

Two subjects (N.B.T. and D.P.) exhibited what seemed to be lasting effects of the bimodal experience. One, (N.B.T.) gave results which indicated a change in the ITD value which was judged as producing a centered binaural image. The direction and magnitude of the change were consistent with the learning of a new ITD-to-position correspondence based on experience in the tailored bimodal experiments. Another subject (D.P.) demonstrated a modification in acoustic image lateralization characteristics which was, if anything, the opposite of this. The third subject (H.F.) was apparently not affected by the experimental manipulations.

In all cases, interpretation of the results was made difficult by extraneous factors, viz.,

(1) the restrictions of the response indicator, i.e., the limited number of push-buttons,

(2) the seemingly natural tendency of subjects to attempt to use all of the push-buttons , even when this may not have been appropriate,

(3) the inconvenient form of the judgement curve (i.e., non-linear) of even the normal acoustic image lateralization results for some listeners (e.g., N.B.T. and H.F.) which made changes in such curves difficult

to assess.

(4) the complicating influence of preliminary signal presentations which demonstrated left and right maxima and the nominal centre of aural and visual image displacements,

(5) the various aspects of human perceptual behaviour which undoubtedly have a bearing in such experiments as these. A subject who by deduction or prior knowledge anticipates the form of signal manipulations or purpose of these experiments would presumably see the combined positional information in a different light from a subject who is convinced (self or otherwise) of the equivalence of the signals.

It was obvious that a modification of experimental technique was in order. The restriction of the response indicator could be simply remedied by the use of an indicator which allowed greater freedom in the range of judgements. However, the other complicating factors were not amenable to such simple solutions.

A number of factors suggested that the difficulties outlined in (2) and (3) were associated with the aspect of experimental procedure mentioned in (4). It would seem natural that unless otherwise instructed, subjects would associate maxima of image displacement with the extreme push-buttons (to give such instructions in some experiments would be to divulge critical information). The non-linear characteristics of judgement trajectories may be due, in part, to the fact that this procedure of establishing maxima of image displacement encourages the subject to set extreme limits of judgements (these limits may or may not be the extremities of the response indicator). Apart from the obvious psychological effects of these restrictions upon judgements, there is the fact that at the extremes, errors in judgement are permitted in one direction only - inwards. Therefore, in the revised experimental technique, physical or conceptual factors which could cause a restriction, in any way, of the positional judgements were avoided.

3.3.0 A Modification of Experimental Technique

The first step in the improvement of the experimental technique was to provide a new positional judgement indicator. A simple and effective solution was found in a device which translates linear motion to rotational motion. The linear motion was provided by a pointer which moved horizontally on a graduated linear scale and the resulting rotational motion provided the drive to a potentiometer. The number of response categories could be altered simply by changing the numerical scale. This positional judgement indication device is illustrated in Fig. 2.5. For experiments which were manually conducted the subject verbally reported the appropriate numerical judgement category.

The next requirement was for a new method of familiarizing subjects with the range of image movement to be encountered during the experiments. A satisfactory solution to this was found to be to present the subject with the complete range of image positions in the first few minutes of an experiment. These signals were presented in random fashion and were judged in the same way as were the remainder of the experimental signals; however, as the first few judgements may have been in error these were usually not included in the averaged results. This scheme replaced the often misleading procedure of presenting only the maxima of image displacement which were, more often than not, associated with the extreme push-buttons. In this modified scheme, subjects were allowed to set their own scales of image displacement but were encouraged to use similar ranges by the simple expedient of telling them in the preliminary instructions that they would probably find that most image displacements fall in the range of, say, ± 5 units. However, they were told to use as much of the scale as was required by their assessments of image displacements.

Experiments conducted using this modified technique were characterized by positional judgement versus ITD curves which were more generally linear in form than those of previous experiments.

3.4.0 Bimodal Experiments: Series 2

The aim of these experiments was specifically to attempt to reproduce the apparent shift of acoustic centre which was exhibited by the subject N.B.T. in the Series 1 experiments. However, on this occasion the basic method was altered.

It appeared from the earlier series that the visual signal was normally dominant, and that the changes which occurred were confined to the characteristics of acoustic image lateralization. On the previous occasions the changes were only clearly observed in experiments using aural signals with different overall ranges of ITD (Figs. 3.2a and b) - a fact which tended to confuse the issue. In the present sequence of experiments therefore the modifications were applied to the visual signal, with the aural signal unmodified throughout.

An additional matter of some interest was whether a simultaneous bimodal presentation was required in order for the cross-modal effects to be elicited. To cast some light on this, the series of experiments included bimodal presentations in which the aural and visual signals were applied and judged sequentially.

In one case the aural signal preceded the visual signal and in the other the sequence was reversed.

These experiments were conducted manually, using the arrangement of equipment described in Section 2.2.0 and illustrated in Fig. 2.2.

The first experiment in the series was a straightforward acoustic image lateralization experiment. The resulting judgement curve which is illustrated in Fig. 3.6(b), E25, is seen to be closely linear, crossing the subjective centre axis at about ITD = -20 µsec.

The next experiment was bimodal, with the signals presented simultaneously. The visual signal was arranged so that it indicated image positions which were judged to be consistently about one judgement division to the right of the acoustic image as lateralized in the preceding experiment. The result (Fig. 3.6b, E26) indicates, as in previous bimodal experiments, that the judgements were dominated by the visual positional information. The subject was apparently not aware of the disparity in positional information.

The experiment which followed was intended to explore the effects of sequential bimodal presentations. In this experiment the aural signal was presented, judged by the subject and then removed; within a fraction of a second the visual image was presented, judged by the subject and then removed. The result for the acoustic image lateralizations (E27A) is interesting. The judged trajectory follows closely that of the previous experiment (E25) over the extreme left-hand region, deviates slightly near ITD = -0.1 msec to almost cross the centre axis, and then crosses at ITD = 0. Over the remainder of the range, the trajectory follows a median course between the curves of E25 and E26. The result for the visual image localizations (E27V) is equally interesting as it suggests that positional judgements pertaining to the antecedent acoustic image had some effect on visual image positional judgements.

Perhaps most surprising of all was the result of the acoustic image lateralization experiment which followed immediately (Fig. 3.6b, E28). This seemingly paradoxical displacement of the judged image trajectory was highly reminiscent of an earlier result (Fig. 3.6a, E24) which also exhibited an unexpected form of displacement.

Continuing the series, a bimodal experiment was conducted using aural and visual signals presented sequentially in the reverse order. On this occasion the visual image localization judgements (E29V) which anteceded the acoustic image lateralization judgements (E29A) seemed to be the dominant factor, but to a lesser extent than when the signals were simultaneously presented (E26).

The final test was an aural experiment, the result of which (Fig. 3.7a, E30) shows that another alteration of acoustic image lateralization performance has occurred. This time, however, the shift was such as to closely align the judged image trajectory with one of the immediately preceding bimodal experiment (E29A).

In summary, Fig. 3.7b illustrates the acoustic image lateralization results (unimodal only) which were obtained at intervals throughout this series of experiments. Included also are results of an experiment conducted prior to,(E24),and one (E31) conducted soon after the sequence just described; these will be seen to closely resemble the first (E25) and last (E30) results of the series, particularly in the important subjective central region. The ITD at which centered acoustic images were reported may be seen to have altered from -20 µsec (E24, E25) to +100 µsec (E28) and then back to -100 and -110 µsec (E30 and E31).

3.4.1 Discussion of Series 2 Experiments

Compared with the results of Series 1 experiments these findings would seem to be much more satisfactory, at least in the sense that there appear to be fewer extraneous factors which could influence the important acoustic image lateralization results. In these experiments the aural signal was a constant factor throughout. Any modifications in lateralization results would appear therefore either to be the consequence of normal random variations, or of influences arising from subjective experience in intervening experiments. In the summarized results illustrated in Fig. 3.7(b), the apparently semi-permanent nature of the modified lateralization characteristics is exemplified by the similar pairs of curves E24 & E25, and E30 & E31. This would seem to reduce the significance of such random effects as may be present.

The shift of judged image centerings from ITD = -20 µsec to ITD = -100 to -110 µsec in Fig. 3.7(b) would seem to be explicable in terms of a newly learned ITD-to-image position correspondence, effected by experience in a bimodal experiment. In other words, a form of adaptation may have taken place. The result of E28 is not so easily described.

3.5.0 Conclusions

A number of effects have been demonstrated which may be attributed to cross-modal influences of signals presented through the aural and visual sense channels either simultaneously or in rapid sequence. Certain of these effects were transitory in nature, viz., the failure of subjects, in appropriate experiments, to detect a disparity in the potential positions of acoustic and visual images, and the judgement, in such circumstances, of a position which was clearly associated with that of the visual image.

Other effects, however, were of a semi-permanent nature. These effects have taken the form of a modification in the relationship a subject demonstrates between the ITD of the binaural signal and the judged acoustic image position. It has been convenient to measure such changes as differences in ITD for which centered images were reported before and after the modification. The largest change demonstrated here was about 0.2 msec (between E28 and E30, Fig. 3.7b).

The magnitude and direction of change have been seen to vary considerably, not only between individuals and occasions, but between different regions of ITD in the same experiment. It is clear that such changes were the results of highly complex interactive effects, and although the underlying causes were not immediately apparent it was of considerable interest that they should have occurred at all. The fact that, in effect, the subjective acoustic centre has been altered may carry some important implications as to the processes of acoustic image lateralizations.

It is therefore clear that specialized knowledge relating to the mechanisms which underlie acoustic image lateralizations would be of fundamental importance to a further extension of this work. The following studies of auditory combinations were initiated partially to provide such information.

On the other hand, the synthesized perceptions which arise when acoustic signals are applied simultaneously to two (left and right ears) channels are of more fundamental importance still than the cross-modal effects which have been demonstrated. It is quite clear that complex psychological and physiological processes are involved in these combinative phenomena. It is proper to enquire if engineering techniques of the kind utilized above could be employed to elucidate at a more fundamental, i.e. physiological, level the mechanisms involved. It is clearly appropriate to restrict studies to two auditory sense channels for this purpose since there is at least some possibility of associating the engineering inferences in some appreciable detail, with current knowledge of electrical neurophysiological activity in the peripheral auditory mechanism.

FIGURE 3.1 (a)

SUBJECT: N.B.T.

E1: Aural
Signal Normal

E2: Aural
Signal Normal

E3: Visual
Signal Normal

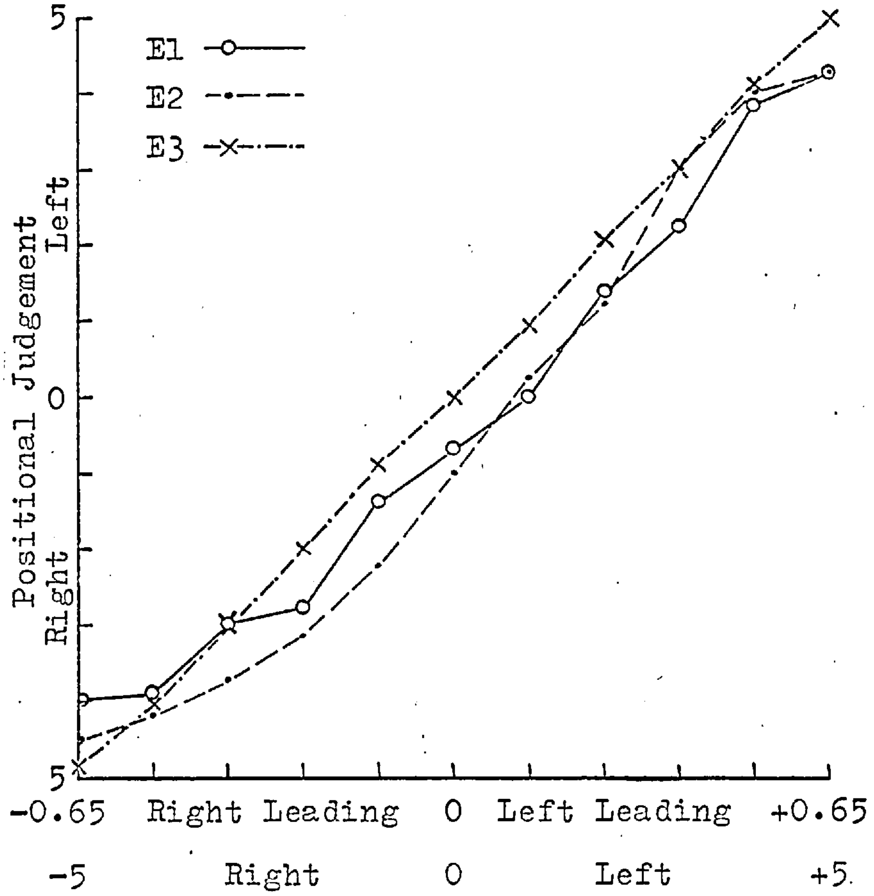


FIGURE 3.1 (b)

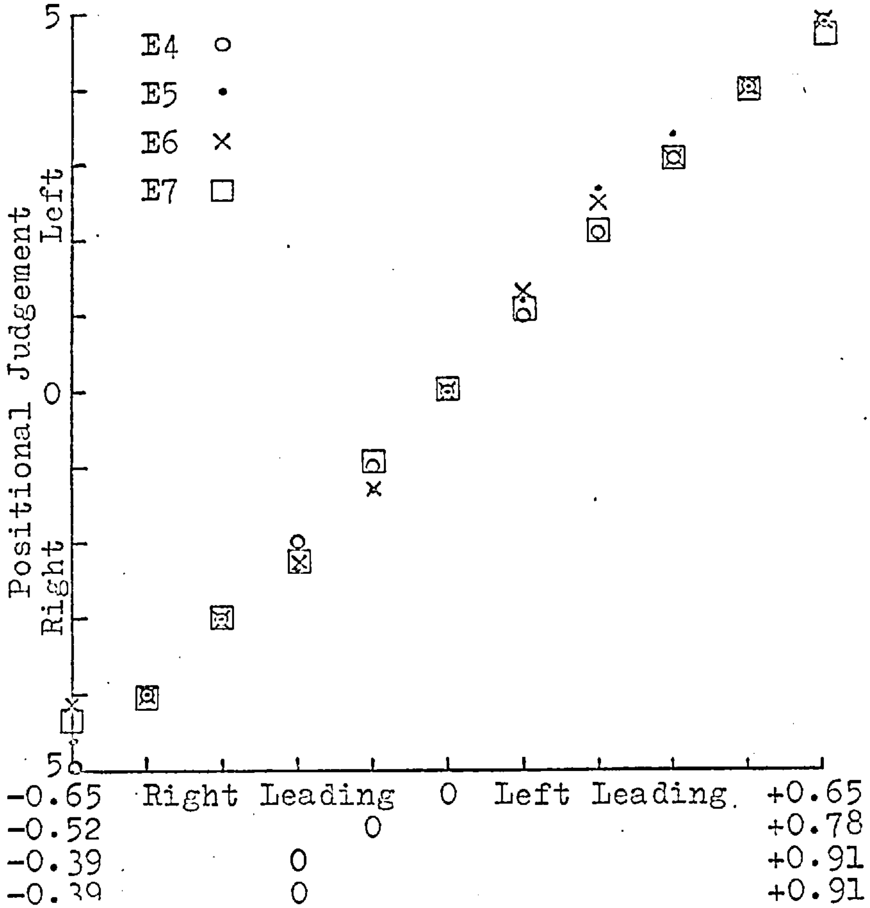
SUBJECT: N.B.T.

E4: Aural
Signal Normal
Visual
Signal Normal

E5: Aural
Signal ITD Bias
Visual
Signal Normal

E6: Aural
Signal ITD Bias
Visual
Signal Normal

E7: Aural
Signal ITD Bias



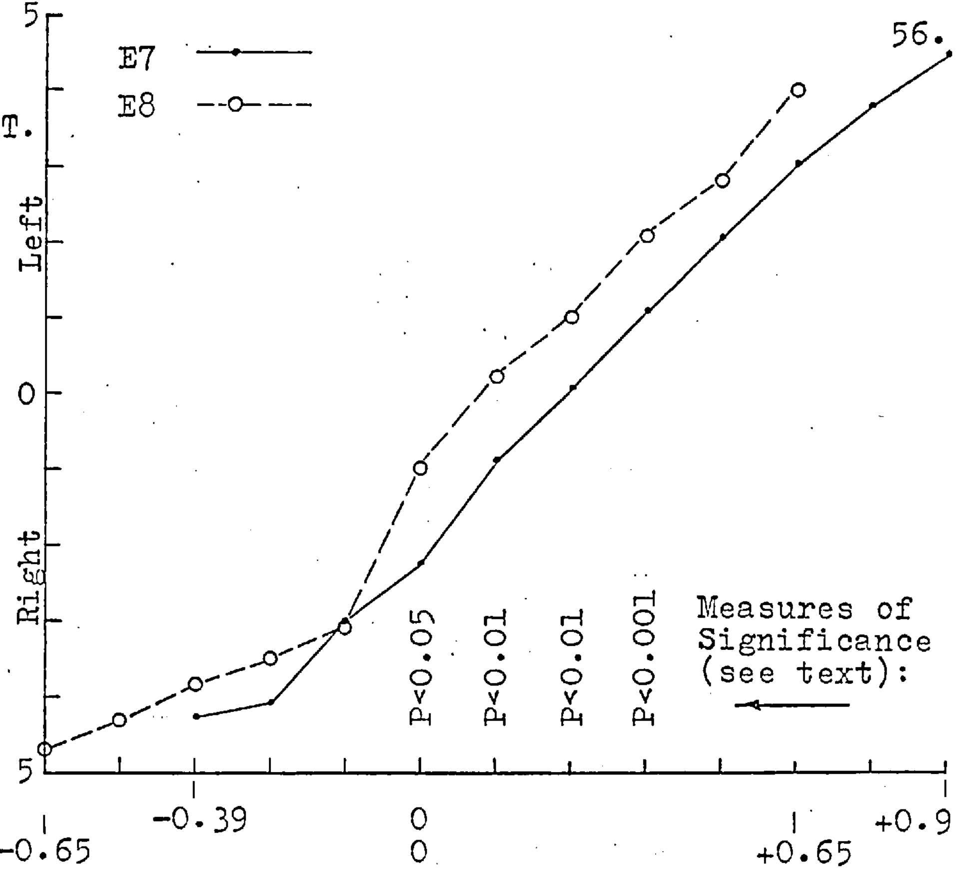


FIG. 3.2(a)

SUBJECT: N.B.T.

E8: Aural
Signal Normal

E7: Aural
Sig. ITD Bias

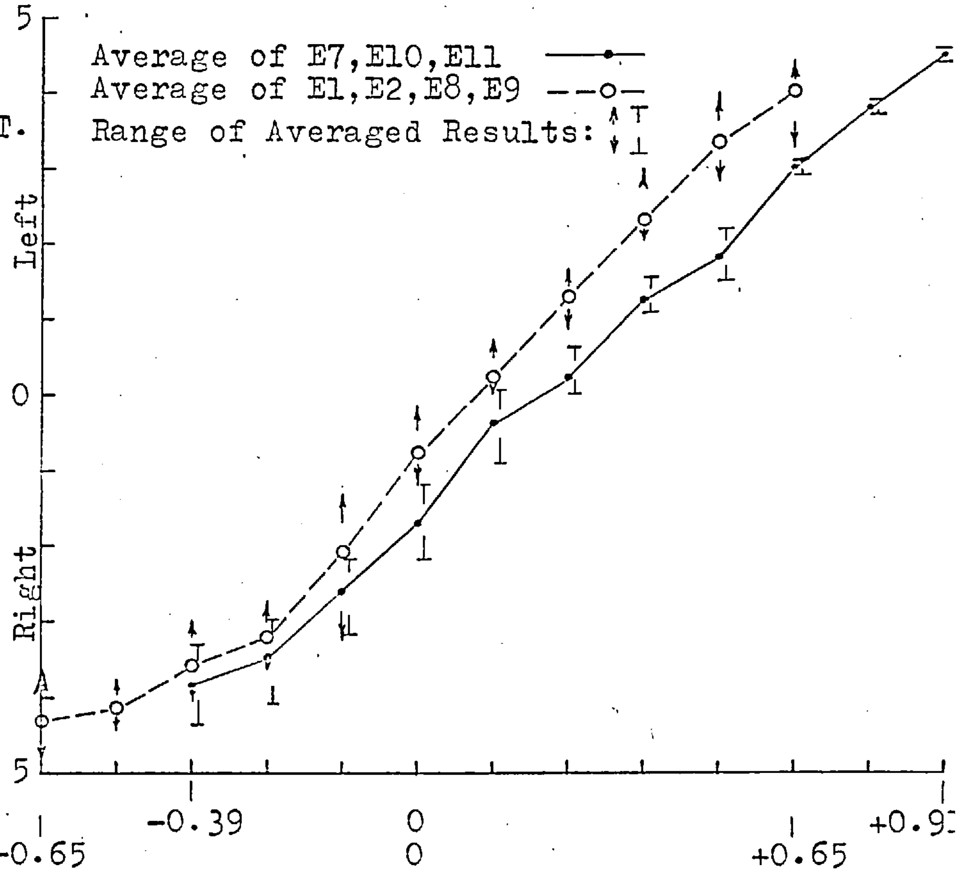


FIG. 3.2(b)

SUBJECT: N.B.T.

E9: Aural
Signal Normal
One Day Later

E10: Aural
Sig. ITD Bias
One Day Later

E11: Aural
Sig. ITD Bias
One Month
Later

FIGURE 3.3(a)

SUBJECT: H.F.

E12: Aural Signal Normal

E13: Visual Signal Normal

E1 and E2: From N.B.T. Figure 3.1(a)

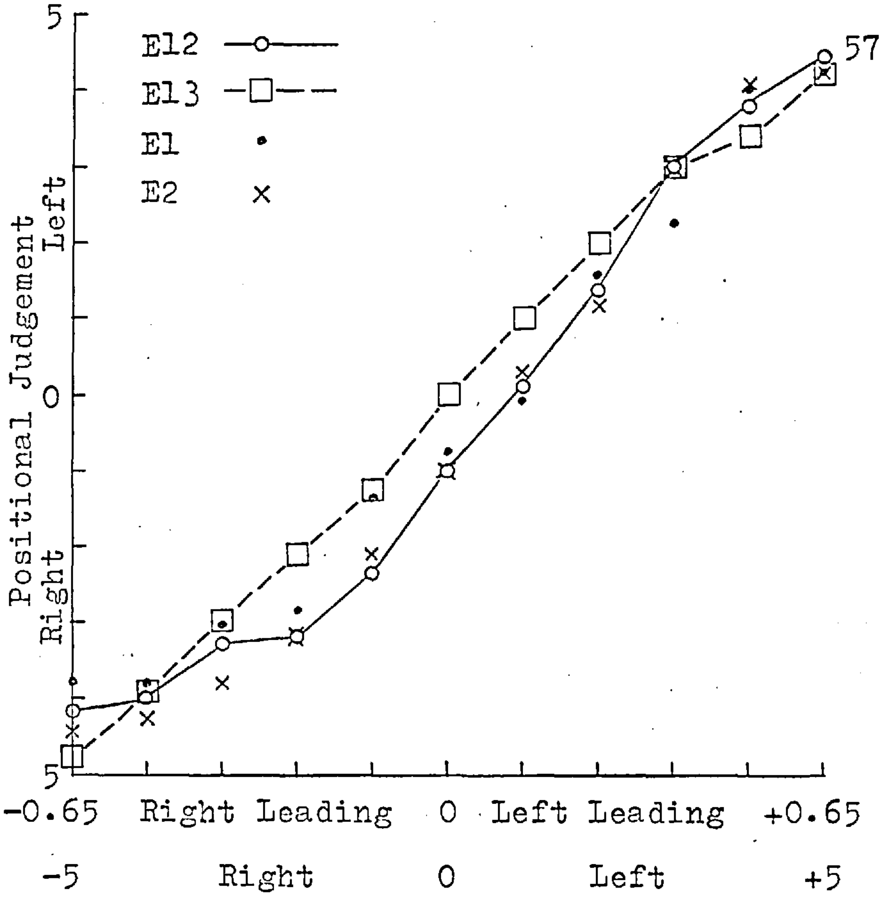


FIGURE 3.3(b)

SUBJECT: H.F.

E14: Aural Signal Normal Visual Signal Normal

E15: Aural Signal ITD Bias Visual Signal Normal

E16: Aural Signal ITD Bias Visual Signal Normal

E17: Aural Signal ITD Bias

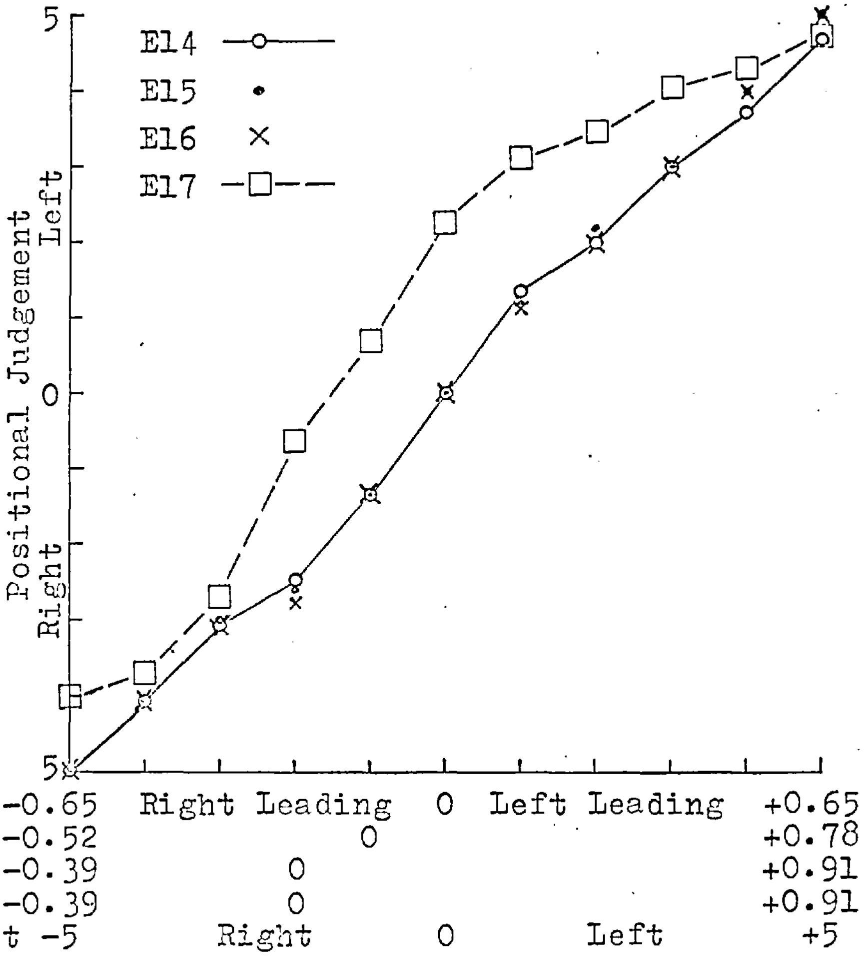


FIG. 3.4(a)

SUBJECT: H.F.

E12: Aural Signal Normal (Fig. 3.3a)

E17: Aural Sig. ITD Bias (Fig. 3.3b)

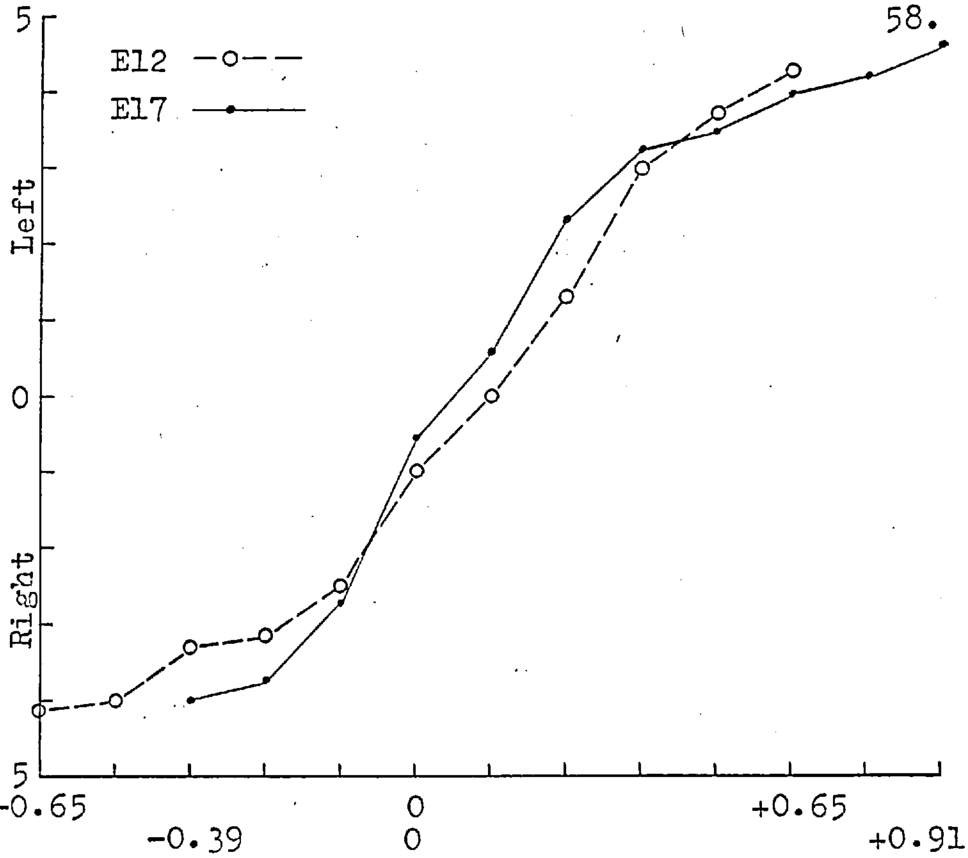


FIG. 3.4(b)

SUBJECT: H.F.

E17: Aural Sig. ITD Bias (Fig. 3.3b)

E7: Aural Sig. ITD Bias (Fig. 3.2a)

E10: Aural Sig. ITD Bias (Fig. 3.2b)

E11: Aural Sig. ITD Bias (Fig. 3.2b)

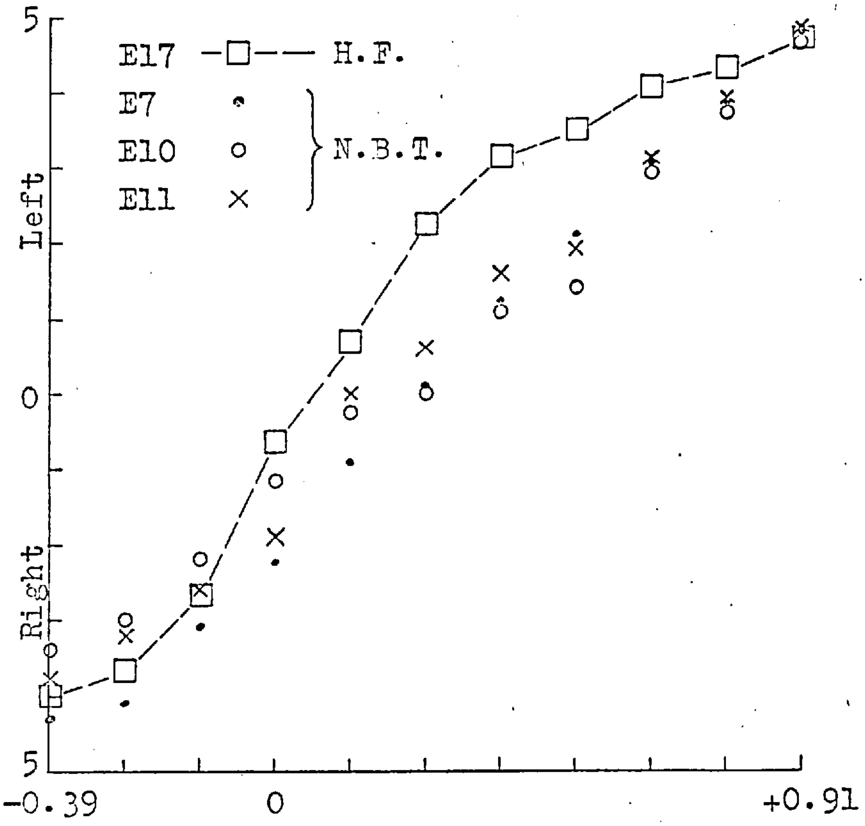


FIGURE 3.5(a)

SUBJECT: D.P.

E18: Aural
Signal Normal

E19: Aural
Signal ITD Bias

E20: Aural
Signal ITD Bias

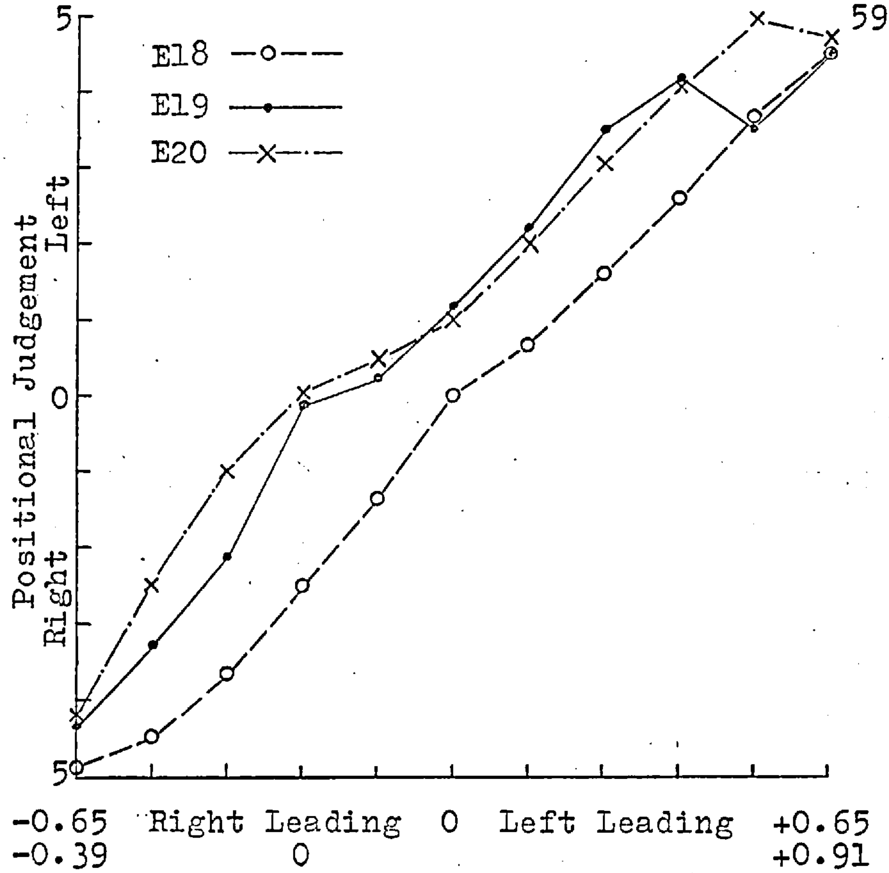


FIGURE 3.5(b)

SUBJECT: D.P.

E21: Aural
Signal ITD Bias
Visual
Signal Normal

E22: Visual
Signal Normal

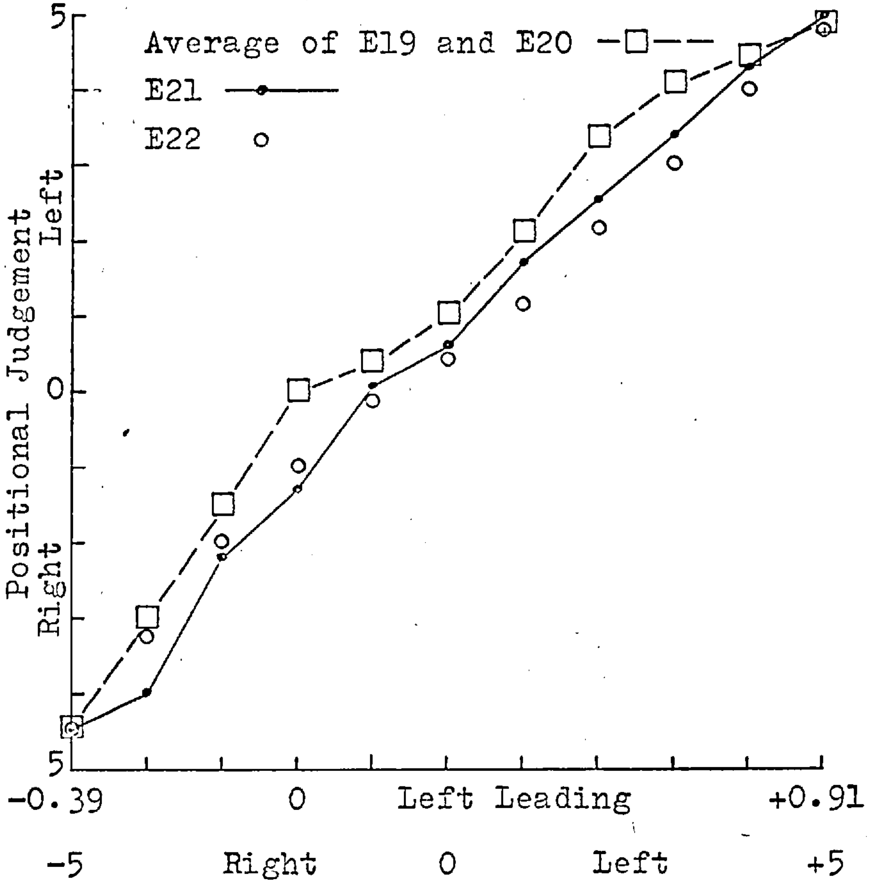


FIGURE 3.6(a)

SUBJECT: D.P.

E18: Aural Signal Normal (Fig. 3.5a)

E19: Aural Signal ITD Bias (Fig. 3.5a)

E20: Aural Signal ITD Bias (Fig. 3.5a)

E23: Aural Signal ITD Bias

Aural Signal: ITD (msec)

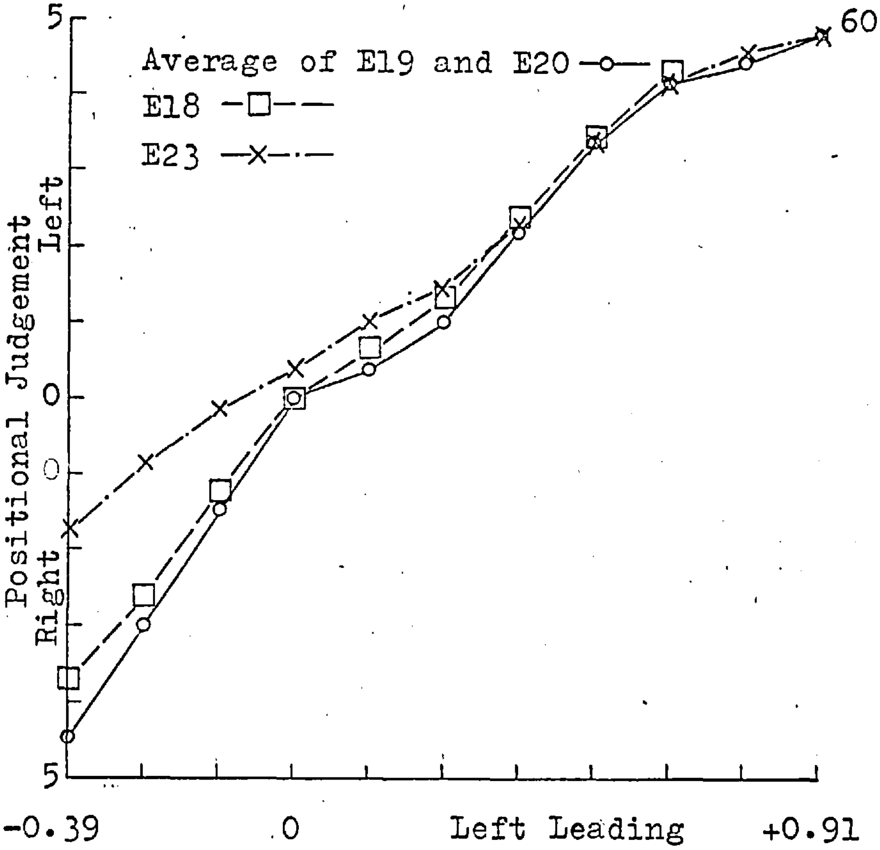


FIGURE 3.6(b)

SUBJECT: N.B.T.

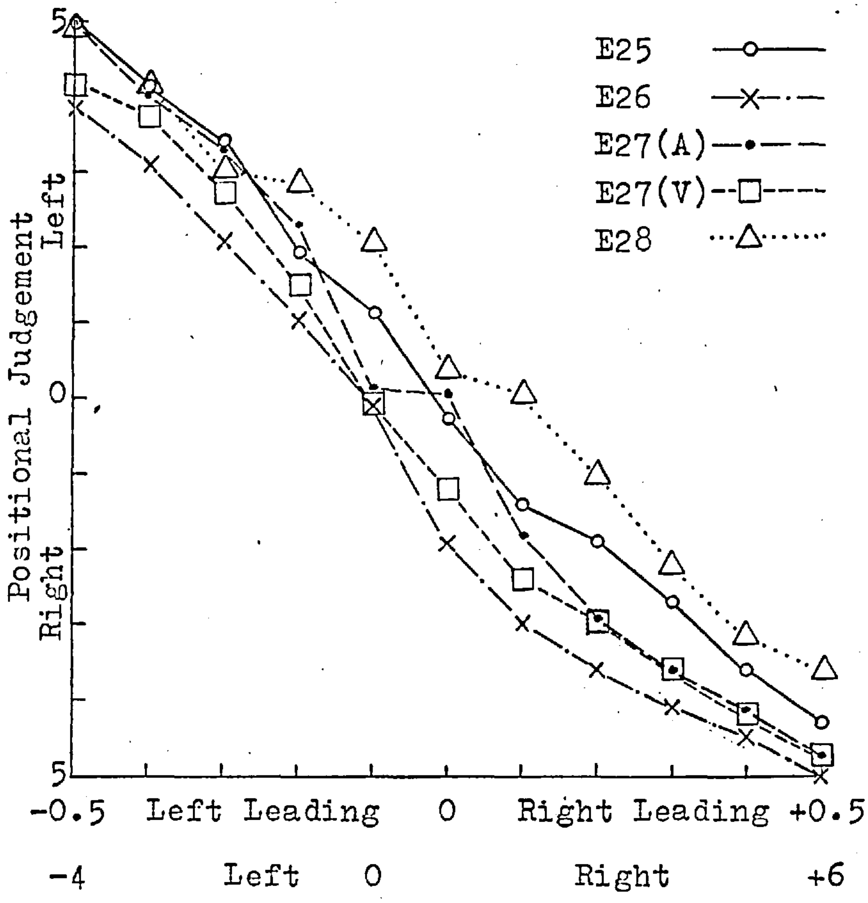
E25: Aural Signal Normal

E26: Aural Signal Normal Visual Signal Bias, Simultaneous Presentation.

E27: Aural Signal Normal Visual Signal Bias, Sequential (A→V) Presentation.

E28: Aural Signal Normal

Aural Signal: ITD (msec)
Visual: X-Axis Displacement (cm)



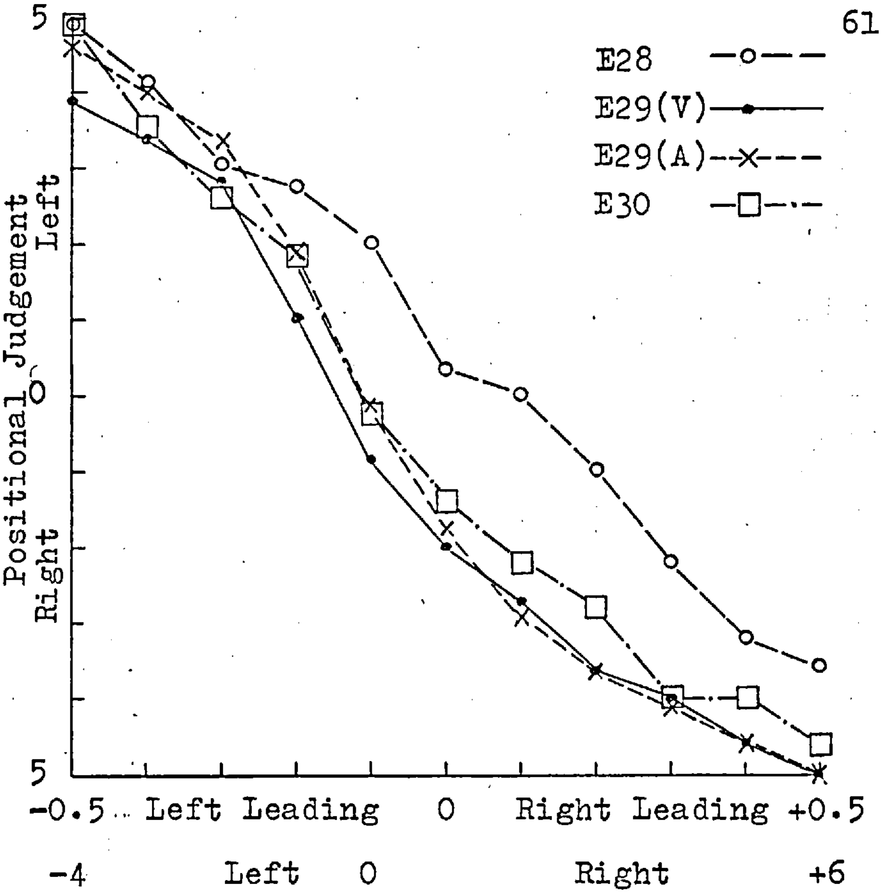


FIGURE 3.7(a)

SUBJECT: N.B.T.

E28: Aural Signal Normal

E29: Aural Signal Normal
Visual Signal Bias
Sequential (V A) Presentation

E30: Aural Signal Normal

Aural Signal: ITD (msec)
Visual: X-Axis Displacement (cm)

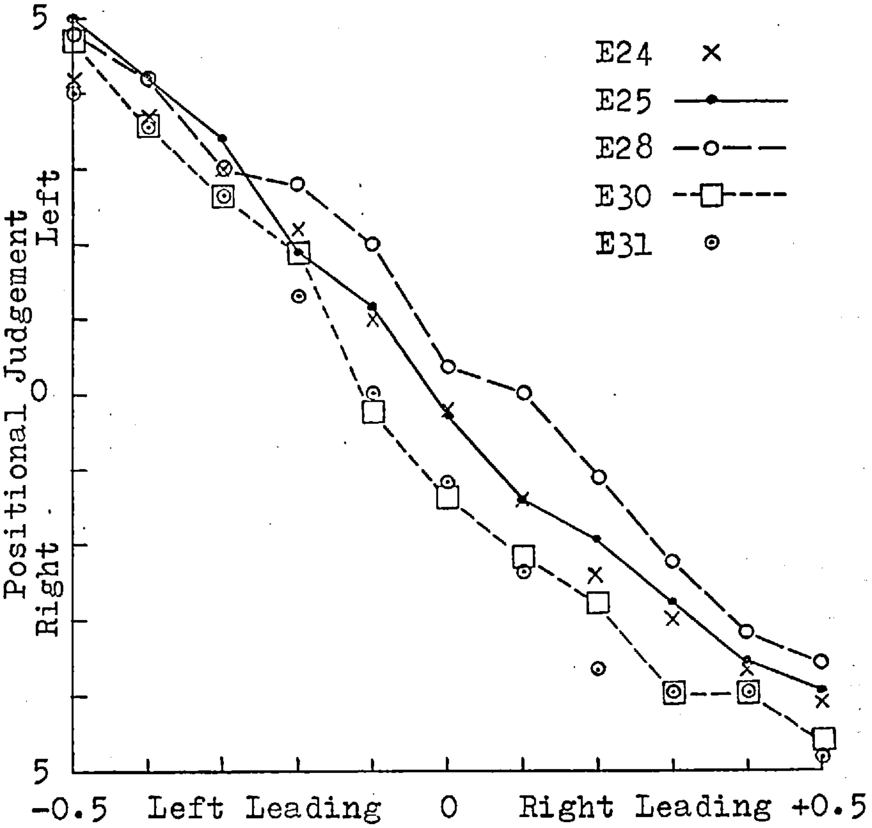


FIGURE 3.7(b)

SUBJECT: N.B.T.

E24: Aural Signal Normal

E25: Aural Signal Normal (Fig. 3.6b)

E28: Aural Signal Normal (Fig. 3.7a)

E30: Aural Signal Normal (Fig. 3.7a)

E31: Aural Signal Normal

Aural Signal: ITD (msec)

CHAPTER 4

A STUDY OF ACOUSTIC IMAGE LATERALIZATIONS

4.1.0 Introduction

Bimodal experiments of the preceding chapter produced some results having important implications regarding the mechanisms underlying acoustic image lateralization judgements. Although it was tempting to speculate about the causes of such findings, it was considered that, as yet, insufficient information relating to the properties of image lateralization judgements in less complex experiments was available. For this reason the present study was undertaken.

It was clear that the previously applied technique of binaural acoustic image lateralization was likely to be instructive in experiments of the required kind. In these experiments the only simultaneous presentation of information was effected through two auditory pathways, i.e., the left and right ears. One general aim of this study was to seek further understanding of the processes which determine the nature of binaural acoustic images and those which underlie lateralization judgements. An ultimate goal, simply related to the purposes of subsequent bimodal studies, would be knowledge which would enable a close prediction of lateralization responses to specific acoustic signals.

But there is an even more powerful reason for exploring in more detail the nature of the combination from left and right auditory channels. It is believed that fusion is a gross representation of a detailed phenomenon of great complexity, in that many complex effects may be shown to simultaneously occur; for example, see Harris (1960), Deatherage (1961), Flanagan (1962) and Sayers (1964). Consequently, it is not now possible simply to discuss the combinational effects of binaural signals by invoking a simple fusion mechanism. This would appear to be particularly important in the circumstances of the current experiments by virtue of the impulsive (i.e., wideband) nature of the auditory signals which were employed in the bimodal signal experiments.

While the concept of fusion, and the logical mechanisms which have been proposed to describe the operation, have been useful in providing a framework for the description of binaural perceptual effects, there was a case to put the mechanism of "fusion" on a more realistic mechanical and physiological basis since only in this way ultimately will understanding most reliable predictions be achieved.

A certain amount of neurophysiological evidence is now available and it has been a significant purpose of this investigation to accommodate this to parallel successfully the main facets of a listener's perceptual experience when receiving information through these two sensory pathways.

The use of repetitive binaural transients (to which attention has been directed due to previous experiments) in this context turns out to be an exceedingly convenient choice for the study of these effects. One interesting finding of some previous work (Sayers and Cherry, 1957)has been the demonstration that considerable advantage accrues from using an experimental signal with a carefully tailored degree of complexity. In that work the transition from a single sinusoid to a two-component tone as an experimental signal turned out to be exceedingly productive in elucidating the salient phenomena. Repetitive binaural transients of the kind considered here may well correspondingly represent an interesting and convenient elaboration from simpler signals.

A large amount of experimental data has been published relating to various properties of binaurally established and perceived acoustic images, whether these are judged to arise naturally in space, or are perceived intracranially, as in earphone listening. Much of this information related to the investigation of factors which may alter the judged spatial position of sound images. The main factors are, of course, interaural time difference (ITD) and interaural amplitude difference (IAD). It was observed at an early stage that, for most signals of impulsive nature, and for low-frequency tones (frequency less than about 1500 cps) the ITD parameter was the dominant factor, although IAD also played an important role. In fact, it was found that to some measure these parameters could be adjusted in such a way that the effect of one off-set the effect of the other. As a consequence many workers studied the quantitative relationships between ITD and IAD and such time -versus - intensity relations have been measured using a variety of signals (see, for example, von Békésy, 1960; Harris, 1960; David, et al., 1959; Moushegian and Jeffress, 1959).

These investigations used several experimental methods, mainly of the type wherein listeners were required to adjust the IAD or ITD factor to bring to the apparent central position the various images which could be identified. Another common method was that of 'matching' wherein listeners adjusted the position of a 'pointer' image to match that of a 'target' image. A third technique was that wherein listeners reported dichotomous 'left' or 'right' judgements indicating to which side of subjective centre the acoustic image was judged to be positioned. There are, however, other methods and other purposes in the study of binaural sound images.

More instructive, in the sense important here, would be data relating subjectively-assessed acoustic image position (i.e., judged extent of image displacement) with the parameters ITD and IAD. Unfortunately, there is little quantitative experimental evidence of this type. Von Békésy (1960) reported subjective impressions of image movements for various signals, but failed to relate these to precise values of ITD or IAD. Guttman (1962) and Teas (1962) each noted the extent of perceived image movement with ITD, but the studies were imprecise, very limited in scope and employed different experimental methods.

Sayers (1964) provided a background of useful information relating listeners' image position judgements to IAD and ITD parameters, in experiments using binaural single and multiple tones. The experimental procedure used in this investigation was basically that which was described in an earlier chapter in connection with the bimodal experiments.

It was clear from the results of Sayers (1964) that the technique of requiring the listeners to report judgements of apparent image position gave experimental results which seemed to be amenable to interpretation and, more important, which were reproducible to an important extent. Moreover, the results gave indications of some characteristics of positional judgement which were speculated may be of primary importance in the study of the image phenomena. One such observation was that for ITD values in the vicinity of T/2 (T=signal repetition period) some listeners reported hearing two images of similar character, one at either extreme position; others, not able to report such a dichotomy, demonstrated an entirely similar grouping of their positional judgements. Equally interesting, however, was the observation of judgements grouped in the central region, for the same ITD values as before. These reports might be interpreted as 'compromise' judgements, representing what could be considered as an apparent balance of attention between two similar-sounding, but spatially separated images.

The influence of IAD on positional judgements was interesting in a not dissimilar way. As well as a general displacement of the judgement curve towards the side of increased loudness, there could be seen an effect which is well described as an earlier transition of attention to the contralateral image as ITD was increased (right-side leading) from zero to T/2.

The essence of the matter could be summarized in a few propositions: first, that simultaneous binaurally created sound images seemed to coexist, and that judgements about the extent of the perceived lateral displacement with ITD could be made reliably over a full cycle of a repetitive binaural signal; second, that listeners were able to attend in some circumstances to either of these; third, that a listener's behaviour was well described by hypothesizing transitions of attention between these images - expressed perhaps by judgements representing an alternation between or a compromise between two separately positioned images.

Such results as those of Sayers clearly formed the basis for a much more detailed study of acoustic images, and it was on this basis that the present study was undertaken. The experiments which follow utilize binaural acoustic transients such as were used in the earlier aural-lateralization experiments. They were chosen primarily because it was argued that such a signal would provide rather more scope for experimental variations than do sinusoids. Sinusoids are difficult to use in perceptual experiments, carry little information, and are not closely representative of normal acoustic signals. Repetitive transients of relatively simple waveform have the particular advantage of being wideband signals not as intractable analytically as speech or noise.

4.2.0 Experimental Technique

The technique is basically that which was employed in the acoustic signal presentations of the bimodal experiments. Listeners were required to report a judgement of the perceived click-or impulsive-image position at each presentation of the binaural signal, and the averaged positional judgements were graphed as a function of ITD. Listeners did not hear changes being made and each value of ITD was used in randomized sequence 4-8 times. Each signal presentation was continued until a result was achieved, requiring a typical duration of about 4-5 seconds. Many listeners found the presence of the graduated visible scale to be of assistance in the quantification of judgements; however, certain of the more experienced subjects preferred to operate without it. In the latter cases no significant differences were observed in the results of experiments conducted with and without the scale.

Listeners were presented with trains of acoustic transients from earphones energized by pulses of 0.1 msec width. Both periodic single-and double-pulse trains were used in various combinations; members of the pulse-pair in the left-ear channel were designated A and D, and those in the right-ear channel, B and C. Single pulse trains would thus comprise only A and B pulses, for the left-and right-ear channels respectively.

Signal levels have been specified relative to the monaural threshold determined using periodic single transients at T = 10 msec. Unit level on the diagrams, therefore, represents the same magnitude as for the individual pulses in these single-pulse trains at about 50 dB SL (levels were referred to the electrical input pulses). Positive polarity indicates an initial condensation peak in the acoustic transient, and negative, an initial rarefaction peak.

Four of the eight (seven male, one female) listeners used in these experiments were experienced in acoustic tests of a similar kind. No formal training procedure was necessary for the naive subjects as none had any difficulty in learning to perform the set task. Prior to each experiment all listeners were reviewed on the details of the task, and presented with randomized signals covering the experimental range of ITD to familiarize them with the experimental environment. When the listener reported his readiness the experiment was started. It was considered desirable to complete the experiment at one sitting - typically about 20-45 minutes depending mainly on the speed of reporting. If, however, the listener reported fatigue or annoyance, the experiment was interrupted and continued after an interval. In such cases extra precaution was taken in the way of further judgements so as to equally represent the possibly modified performance after the interval, although generally such changes were not observed. To gain most benefit from these very tiring experiments the experimenter would often devote extra attention to those values of ITD which occasioned widely varying positional judgements at different presentations, after having first covered the full range of ITD values a requisite minimum number of four times.

4.3.0 Apparatus

The equipment used to produce the acoustic signal for the bimodal experiments was modified to accommodate a range of ITD covering at least one period of the repetitive transient signal, and to allow for independent manipulation of period (T), pulse amplitude, and number of pulses per period. This was accomplished with the system illustrated in Fig. 4.1.

The period of the repetitive signal was set and maintained by a 'clock' pulse generator, which served also as the first variable-delay generator (Delay 1); this delayed output was used to trigger a second variable-delay generator (Delay 2). The output from the Delay 2 generator triggered the signal-pulse generator for the left-channel and the non-delayed output from the clock generator was used to excite two signal-pulse generators, the summed outputs of which provided the right-channel signal. The output of each signal-pulse generator was passed through an attenuator and a gated summing amplifier to the earphone power amplifier.

The signal-pulse generators were each capable of producing uniform trains of single-or double-pulses, each pulse 0.1 msec duration, with continuously variable pulse-pair separations. Repetition period, T and pulse-pair spacing δt, were set and monitored using a digital frequency meter and interval counter.

The ITD was established by the cumulative delays introduced by the Delay 1 and Delay 2 generators. The use of two of these units enabled ITD values covering one complete period of the signal to be obtained. The experimenter set the ITD by manipulating the two delay generators while observing the relative temporal spacings of the left-and right-earphone signals as displayed on an oscilloscope (Tektronix 545). It was found convenient to use the 0.2 cm markings on the calibrated graticule as the ITD intervals in most experiments, at the usual time-base setting of 1 msec/cm each interval corresponded to 200 μsec. In certain experiments larger intervals were sometimes used in part, or throughout. This was done only if the particular experiments would otherwise have been of unreasonable duration, and if it was clear that no important information would be lost by such a modification of technique.

Listeners were again seated in the sound-treated booth, and communicated with the experimenter by means of an intercom system.

The earphones used in these experiments were Sharpe HA-10 units, matched and calibrated by the manufacturer to within ±2 dB from 125 cps - 3kc, and ± 3 dB from 3kc-8kc. A number of different earphones was tried in the experiments. The acoustic output was investigated in each case, using an artificial ear (type 4109 by Bruel and Kjaer) with a flat-plate coupler for circumaural earphones (see Shaw, 1962), and a 6 cc. coupler for the standard-type earphones. The earphones used were STC type 4026 A, Permoflux PDR-10, S.G. Brown Super K, Sharpe type HA-10, and units comprising TDH-39 inserts in Anticounstic Company surrounds. The latter were especially developed for use in the psychoacoustics group of this department. The differences in experimental results obtained using these various units, though noticeable, were much less significant than any of the effects discussed here. It is to be noted that such effects may be important in some experiments, e.g., some in Chapters

6 and 7, but the situations would not be as below. Acoustic responses are illustrated in Fig. 4.2. The ultimate choice of the Sharpe HA-10 was based on a number of factors: high calibre of acoustic performance, effective attentuation of ambient noise (approximately 40 dB at 1 kc/s), universal availability and comfortable fit. The last point can be very important in prolonged experiments. It was found that uncomfortable earphones contributed substantially to listener fatigue. Moreover, this type provides adequate space within the circumaural surround to accommodate the pinna, distortion of which could result in pressure on the auditory canal and possible impairment of aural function.

4.4.0 Experiments

The following experiments were planned to elucidate the character and origins of the dominant sound image which was perceived by listeners presented with binaural trains of single and double acoustic transients. Judged image trajectories with ITD were determined in experiments using a variety of different repetitive patterns of acoustic transients applied at equal level as well as under the influence of interaural amplitude disparity.

4.4.1 Trajectories of Judged Image Position with ITD: IAD = 0

The curve of Fig. 4.3 illustrates the trajectory of judged image position for the simple case of binaural periodic single transients (A and B pulses only, T = 6.0 msec, IAD = 0). For ITD increasing monotonically the image is observed to be tracked across the auditory space, crossing centre at closely ITD = 0, and resting at the lateral extreme position until the image which is hypothesized to appear at the contralateral extreme position becomes of comparable importance (at ITD = approx. T/2). Judgements then swing towards and are finally dominated by its apparent spatial position.

With periodic double pulses applied to the right earphone each pulse (A) in the left-ear channel is then associated with a pulse pair (B and C) in the right-ear channel. The curves of Figs. 4.4 - 4.8 show, for one listener, the influence of reducing the spacing (δt) between pulses B and C of the pair while maintaining the repetition period constant at 6 msec; all pulses were of equal level (A:B:C = 1:1:1).

Figure 4.4 shows what could be interpreted as the presence of two images; one, S_1, arising from binaural interaction of A and B transients and the second, S_2, arising from binaural interaction of A and C transients. Here, where δt was set equal to one-half the repetition period the judged trajectories of the S_1 and S_2 images were identical; however, in Figs. 4.5 - 4.8 the S_2 image can be seen to be progressively less well represented in the positional judgements as δt was reduced.

Figures 4.9, 4.10, 4.11 illustrate for the same listener, results of experiments in which the $\delta t/T$ proportion for the experiments of Figs. 4.4, 4.7 and 4.8 were reproduced at a repetition period of 10 msec. Results of similar experiments using two other listeners are illustrated by the curves of Figs. 4.12, 4.13, 4.14 and Figs. 4.15, 4.16 and 4.17. At this lower repetition rate the important features in the shapes of the curves are generally similar to those observed in comparable experiments at the higher repetition rate, particularly for those results obtained using the same listener. It is clear that although the importance of the S_2 image seems to be diminished with a reduction in pulse-pair spacing, in each case the listener appears to track first one image and then the other.

Comparison of results obtained with the three listeners (Figs. 4.10 - 4.17) reveals some interesting points of difference, viz., the choice of different numerical scales in the quantification of judgements, the different forms of transition (e.g., the rapid $S_2 - S_1$ transitions of F.E.T. and R.E.T. versus the gradual transitions of J.P.B.), and the apparently complex situation implied in the curve of Fig. 4.16 versus the relatively simple situations suggested in the graphs of Figs. 4.10 and 4.13. The latter contrast is particularly interesting, for it shows that while there may be other, as yet unidentified, factors influencing the performance of some listeners, certain features of the results are generally similar, namely, the basic S_1 and S_2 image trajectory pattern.

A point which emerges from a comparison of Figs. 4.9, 4.10, 4.11 with Figs. 4.4, 4.7, 4.8 is that the shapes of the curves would appear to be more dependent on pulse spacing relative to repetition period than upon absolute pulse spacing. Because there is an apparent limit to the lateral displacement of an acoustic image, the shape of curves associated with large repetition periods may be somewhat modified at the extremes compared to those obtained at smaller repetition periods.

The curves of Fig. 4.18 illustrate the results for binaural double-pulse trains for two listeners and for 6 and 10 msec repetition periods, ${}^{\delta t}/T = 0.3$ for both signals and both listeners. Results are plotted against ITD as percent of period T. In the curves there is evidence to suggest the existence of three image trajectories per period. The dominant image trajectory appears to be that expected for the hypothesized positionally-coincident images formed by the binaural interaction of Pulse A with B and D with C. To the left and right of this main image crossing are indications of images due possibly to the hypothesized fusions of pulses D with B, and A with C.

4.4.2 Trajectories of Judged Image Position with ITD: Influence of Relative Level

Again beginning with the simple case of binaural periodic single transients, the curves of Figs. 4.19, 4.20, 4.21 illustrate for two listeners judged image trajectories for T = 6 msec and T = 3 msec. In these experiments fixed magnitudes of lAD were introduced that, by steps, increasingly favoured one side. Complete sets of judgements were taken at each IAD. The figures show the resulting families of curves superimposed for purposes of comparison. Note the similarity between these curves and those found using binaural single tones (Fig. 4, Sayers, 1964). The degree of lateral bias of judgements with IAD is not the same for all listeners, as is evident from a contrast of Figs. 4.20 and 4.21.

Results of experiments using binaural single versus-double pulse trains in which the levels of A, B and C pulses were individually manipulated are shown in Figs. 4.22 - 4.28. The curves of Fig. 4.22 illustrate the effect of altering the level of pulse A alone. The effect is well described as a simple vertical shift of the curve with increasing IAD. Figs. 4.23 and 4.24 illustrate the effect of altering the level of pulse C alone while Figs. 4.25 and 4.26 involve a level change of pulse B alone. The latter figures represent results using two listeners for comparison. It is again clear that both S_1 and S_2 images are present and that changing the relative levels of the B and C pulses merely influences the degree to which the S_1 and S_2 images individually contribute to the overall pattern of positional judgements. It is possible to observe what may be an independent sidedness bias to the S_1 and S_2 image judgements due to the IAD existing between the pulses A and B and the pulses A and C which in combination form these images. A second observation relates to the transition of judgements from following the S_1 image to following the S_2 image, and back again to the S_1 image. Experiments in which either pulse B was reduced or pulse C increased in amplitude, illustrated by Figs. 4.23 - 4.26 indicate progressively earlier transitions from S_1 to S_2 and progressively later transitions from S_2 to S_1. This latter transition, which for IAD = 0 occurred about midway between S_1 centerings, for maximum IAD occurred nearly midway between S_2 centerings.

A special case of these experiments worthy of individual attention is that in which $\delta t = T/2$. The curves of Figs. 4.27 and 4.28 illustrate for this situation the effect of reducing the level of the C pulse; results are shown for two listeners at two values of T.

All listeners under some conditions gave indication of the presence of an unexpected subsidiary image which could be described as following a trajectory spaced approximately 1 msec from a major image trajectory. Often its existence could be inferred only from a slight indentation of the judgement curve, but occasionally listeners would show, very clearly, judgements which transitioned to follow this new image; two examples are illustrated in Figure 4.29. Other examples may be seen in Figs. 4.15 and 4.16. This effect has been less pronounced in experiments employing the shorter repetition periods, although it did not appear to be seriously influenced by the introduction of IAD.

4.5.0 Discussion

Results of the preceding experiments demonstrate that, for the binaural signals used here, listeners are able to judge the apparent position of a sound image as it is perceived in what might be called his "acoustic space". This has been illustrated earlier (Sayers, 1964) for tones, and here for periodic binaural single and double transients with ITD ranging over one repetition period of the signal and for various conditions of IAD.

In the cases of binaural single tones and binaural single repetitive transients it would appear that binaural interaction of the left-and right-ear signals gives rise to at least one sound image which occupies a position in the listener's perceived auditory space according to relative ITD and IAD. For each of these signals as ITD is varied monotonically from zero, the image is observed to move almost linearly towards the ear at which the signal leads in time. Depending on the repetition period of the signal and the prevailing IAD, the image may or may not reach an apparent limiting position at the lateral extreme before the appearance, (at about ITD = T/2), of a second image at the contralateral extreme position apparently causes a division of the listener's attention. In this latter region judgements tend to fall into three groups: the left and the right extreme positions, or some intermediate point which, perhaps, indicates a balance of attention between these two similar-sounding images at the opposite extremes. A number of listeners reported hearing these images simultaneously, and in such cases simply reported the two apparent positions; others, when giving single judgements,

ments, often remarked on the 'dispersed' or 'split' character of the image. Some support of these descriptions was given by the variation of positional judgements in such instances.

In the case of binaural single - versus double-pulse trains one would expect that at least two images would be established, perceived at different positions, according to the separate interaural time relations between pulse A at one ear, and each of the pulses B and C at the contralateral ear. Again, for convenience, the image (AB) formed with the leading member of the pulse pair has been designated S_1 and the image (AC) formed with the trailing member, as S_2.

The results suggest that as ITD is increased from zero, the listener's positional judgements follow first the S_1 image, transitioning at some stage to follow the S_2 image and again, later, to once more follow the S_1 image. This view is demonstrated quite clearly in the Figs. 4.4 through 4.17, where in all cases the listener appears to track first the S_1 image and then the S_2 image. There is, however, in the sequences of experiments illustrated by Figs. 4.4 - 4.8, 4.9 - 4.11, 4.12 - 4.14 and 4.15 - 4.17, a gradation of effect between situations where S_1 and S_2 trajectories are clearly identified and where S_1 appears more-or-less alone. The S_2 image would appear to contribute less to the overall pattern of positional judgements as the pulse spacing δt is reduced.

The implication is that both S_1 and S_2 images are present but that only when S_2 is near the centre of the perceived auditory space does it start to contribute to and then dominate the judgements. As δt is reduced relative to T, there is proportionately a later transition from S_1 to S_2, indicating that S_2 is, in some important sense, of less significance. The argument is illustrated by the idealized curve of Fig. 4.30, depicting the experimental conditions of Fig. 4.5.

A second manipulation which results in a modification of the relative importance to the judgement of the S_1 and S_2 images is one in which the relative amplitudes of the B and C pulses are modified. The curves of Figs. 4.23 - 4.28 illustrate the results of such experiments. It is clear that either a decrease in the level of the B pulse relative to the C pulse, or an increase of the C pulse relative to the B (the unaltered C or B pulse amplitudes, respectively, being equal in level to the contralateral A pulse) results in a diminished importance of the S_1 image. This effect is generally described as an earlier transition of judgement from the S_1 to the S_2 image trajectory, and a later transition from the S_2 to the S_1 image trajectory. Figs. 4.23 and 4.24 demonstrate the manner in which the judged image trajectories alter from a pattern dominated by the S_1 image to one dominated by the S_2 image trajectory as the C pulse is increased in amplitude. Similar trends are to be seen in Figs. 4.25 and 4.26 illustrating, for two listeners, results of experiments in which the level of the transient due to the B pulse has been reduced in amplitude.

Figs. 4.27 and 4.28 illustrate the results of experiments in which the pulse-pair spacing δt was set at T/2 and the level of pulse C reduced. Again, the relative importance of the S_2 image diminished as expected, with transitions from the S_1 trajectory to the following S_2 trajectory occurring later and from S_2 back to S_1, earlier. The argument is demonstrated by the stylized curves of Fig. 4.31. Vertical displacements of the S_2 trajectories which would be expected due to the increasing IAD between the A and C pulses (which together form S_2) have been ignored.

Results of experiments in which equal-level binaural single versus double-pulse trains were employed, suggest that monaural temporal masking effects may be an important factor. This concept is not new: v. Békésy (1930) reported the apparent suppression of the later of two sound images, Wallach, et al. (1949) investigated this "precedence effect" in some detail, and more recently Guttman, et al. (1960) and Harris, et al. (1963) used signals similar to those presently employed to specifically study monaural temporal masking effects and relevant neural mechanisms.

Guttman, et al. (1960) and Harris, et al. (1963) applied the experimental method of 'centering'; listeners in these experiments manipulated the ITD between the contralateral single-and double-pulse trains to bring to the central position any images that could be found. Guttman, et al. found that the image S_2 was "silenced" for δt less than 3.0 msec, even at the highest repetition rate of 125 pps. However, results described in Section 4.1 show very definite evidence of the S_2 image at δt intervals much smaller than this minimum figure, at the lower repetition rate of 100 pps (T = 10.0 msec).

A somewhat different result is reported by Harris, et al. where listeners using the same 'centering' technique were able to centre images arising from pulses spaced by as little as 0.5 msec (i.e., δt = 0.5 msec). A curious finding was that listeners made no significant centerings of the S_2 image when δt = 2 msec. The authors had some difficulty in accounting for this observation, and put forward various proposals (e.g. specialized neural "gates") in part explanation. The image lateralization experiments reported here show that the images are, in fact, unequivocally present, and it again seems that failure to achieve centering is indeed no evidence at all for the absence of the second image.

A further clarification from the experiments is extended to the forward - and backward - masking phenomena discussed by Guttman, et al. (1960). In these experiments the authors have defined forward masking as an "inability to fuse the second click" (the C with A), and backward masking as an "inability to fuse the first click" (B with A). The results of experiments here suggest a rather different interpretation. For example, the curves of Figs. 4.7 and 4.6 exhibit what would be measured by the technique of Guttman, et al. as forward masking, those of Figs. 4.14 and 4.25 (1:0.35:1) what would be measured as backward masking. It seems clear that, in fact, both S_1 and S_2 images are perceived (i.e., fused) and are independently lateralized at different times. The failure of either image to appear in the subjective central region seems to indicate merely that the listener's attention has been diverted from S_1 to S_2 at different ITD. While it seems that masking effects may indeed be involved, the present results demand a somewhat different interpretation.

In common with the findings of Harris, et al. and Guttman,et al, the present results indicate that for a given double-pulse spacing δt the degree to which the S_1 and S_2 images contribute to the overall judgement pattern is a function of period T. In other words the ratio $\delta t/T$ appears (at least over the 10-5 msec range of T investigated) to be the single factor which dominates the judged image trajectories when all pulses are of equal level. Guttman, et al. have proposed a model to account for such an apparent improvement of "temporal resolution."

A further observation to be made from the findings of Harris, et al. (1963) is the occurrence of image centerings spaced about 1 msec behind those of the major S_1 and S_2 images. The authors argued that this effect could be explained on the basis of damped oscillatory vibrations of the basilar membrane. On the assumption that neural firings occur at some threshold on unipolar upward displacement of the basilar membrane (in the direction of the tectorial membrane) and that the most significant neural information is derived from the membrane region moving with the greatest amplitude, one would expect such oscillatory neural firings to occur at about 1 msec intervals ((observation using a computational model (Flanagan, 1962) of the basilar membrane indicated a maximum of displacement in the vicinity of 1000 cps)). Such an explanation seems logical and would appear to fit the facts well.

As was described in a previous section (4.4.2) all listeners, on occasion, gave evidence of the presence of a subsidiary image, This image appeared to follow a trajectory spaced about 1 msec from either the S_1 or S_2 image trajectories; hence, in a different way the finding of Harris, et al. is confirmed.

It was speculated earlier that the simultaneous presence of two similar-sounding, but spatially-separated images could lead to positional judgements associated with either one of these images (or both) or with some intermediate or compromise position. This is a crucial matter for it enables an interpretation of the form of positional judgements in the transitional regions between the apparently independent tracking of the S_1 and S_2 images. It is suggested that single positional judgements under such circumstances will be entirely dominated by or biased towards the image which is of greater importance in some sense. Experiments have shown that this relative importance of the S_1 and S_2 images is apparently modified by a change in level of either the B or the C pulse or by manipulation of the δt. A change of the $\delta t/T$ ratio clearly changes the potential masking of the C by the B pulse and of the B by the preceding C pulse. Level alterations of B and C pulses at constant $\delta t/T$ ratio, would simply be enhanced by the monaural masking process as envisioned by Guttman, et al. (1960). Certain results, for example Figs. 4.27 and 4.28, seem to carry implications as to the significance of the subjective central region in the mechanism whereby a listener selects which image to track. It would appear that the lateralization potential of an image increases as it approaches the central region. In fact, this argument can be applied to most of the reported experiments, bearing in mind the particular prevailing IAD condition.

The concept of image trajectories and the principles of image positional judgement as elucidated thus far have been applied to a variety of experimental findings, and have proved to be of merit in clarifying certain, otherwise ambiguous, results. It is now of interest to apply these methods to a still more complex experimental situation.

Wallach, et al. (1949) in their study of what was called the "precedence effect" used, as a signal, pulse-pairs in each channel, and considered image position as a function of the pulse-pair intervals and ITD. Listeners in these experiments reported image locations along an arc from ear to ear which was divided into 30^{o} sectors. In the analysis these positional judgements were pooled into simple "left" and "right" categories, and the resulting percentage judgements "left" was graphed as a function of the independent parameter. The thresholds as determined by this method of limits were then the important experimental results.

The first finding reported by the authors in their series of experiments was an apparent "loss of precision in distinguishing 'right' from 'left' when the first pair of disparate clicks is followed after 2 msec by a second simultaneous pair,"(see Figure 4.32(a) which reproduces Fig. 7 of Wallach et al.1949). The implication is that when the ITD between the C and D pulses is fixed at zero, and the ITD between the A and B pulses is altered (the AD and BC intervals are approximately 2.0 msec), the resultant sound image is displaced by a smaller amount per unit ITD than that perceived when the A and B pulses are presented alone. Note that in this discussion the A, B C and D pulses referred to are those of the present author's notation (i.e., A and D in one channel, B and C in the other) and are <u>not</u> those of Wallach, et al.

Secondly, when a fixed non-zero ITD is established between the C and D pulses the judgements are biased towards the side at which either the C or D pulse leads in time. The magnitude of the bias is shown to be a function of the fixed ITD (i.e. C-D), increasing with this ITD up to about 400 µsec and then decreasing as the ITD is increased further. This effect was verified in an experiment in which the ITD between A and B pulses was set to zero, and that between D and C was adjusted as the independent variable; see Figure 4.32(b).

The arguments used in describing results of the present experiments when applied to these findings permit possible explanations for the reported phenomena.

Consider first the results of experiments in which the ITD between A and B pulses was the independent variable, with the ITD between C and D pulses fixed at some value; see Fig. 4.32(a). In these experiments conditions are favourable for the establishment of four images: the binaural interactions of A with B, D with C, A with C and D with B, presumably each producing an image. The ITD between pulses forming the latter two images is considerable (2 msec) and these images would therefore be expected to be symmetrically disposed at contralateral extreme positions. Because the ITD between A and B pulses is varied over over only a small (± 100 μsec) range, it may be reasonable to ignore these two images (A-C and B-D) in the discussion as they represent constant and secondary factors.

For fixed ITD = 0 the image arising from an interaction of C and D pulses would be centered, while that due to A and B pulses would have a position determined by the particular ITD relation existing between them (the independent variable, in this case). For ITD values over the range ± 100 μsec the AB image may be expected to move slightly left and right of centre. In this multiple-image situation positional judgements could relate to either the A-B or the C-D images or to some intermediate position. It is, therefore, not surprising that the lateralization thresholds measured for this situation are larger than those measured for the AB image alone. As expected, the 50 percent crossing points are not seriously altered by the presence of the fixed central image.

Where the fixed ITD establishes a C-D image to the left or right of centre, then on the same basis one would expect an overall bias of judgements to the left or right sides respectively. Again this would be attributed to judgements being distributed at and between the apparent positions at the A-B and C-D images. The experimental results confirm this. Experiments in which the A-B image was fixed at centre and the C-D image moved about, naturally gave similar results.

The magnitude of the lateral bias induced in positional judgements by the off-centre C-D image was shown to be a function of the ITD between the C and D pulses. An explanation of this result is assisted by the following assumptions: assume that a 600 μsec ITD is sufficient to displace an acoustic image completely or almost completely to one side (see experiments of this and earlier chapters) and that further increase in ITD serves merely to diminish the importance of this image in the judgement (e.g., see Fig. 4.3 where judgements transition from one side to the other when at ITD = T/2 the contralaterally positioned images are of apparently equivalent importance). Assume also that the subjective central region is an important feature in the judgement mechanism (e.g. see Figs. 4.27 and 4.28 which suggest that of two coexisting images the one closer to the centre position is more likely to dominate the judgements). These assumptions apply also to the results reported in section 4.0 and were, in part, suggested by certain of them.

Consider the case of a centered and fixed A-B image and a roving C-D image. As the independent variable (ITD) is increased from zero the C-D image is displaced away from centre. Judgements may be expected to be distributed at and between centre and the position of the C-D image. It is speculated that at some point, the spatial separation of the two images and the remoteness of the C-D image from centre combine to diminish the importance of the C-D image in the judgements. At some further point the hypothesized devaluation of importance of the C-D image with increasing ITD would contribute to the almost complete domination of judgements by the A-B image.

In short, the lateral bias in the appropriate direction of positional judgements would be expected to increase with increasing ITD between the C and D pulses. The bias would pass a maximum and would begin to diminish as ITD approaches about 600 µsec, falling off still more with larger values of ITD. The actual result (shown in Fig. 4.32(b) which depicts Fig. 13 of Wallach, et al., 1949) would appear to be well described by this interpretation.

A final matter to consider is that of the ITD/IAD trading ratio proposed by certain authors. This measure would seem logical on the basis that an image displacement caused by a change of one of these factors may, to a certain extent, be compensated for by suitable manipulation of the other thus trading an interaural <u>time</u> for an interaural <u>amplitude</u> difference, or vice versa. Some workers involved in such studies have used the method of 'centering' the acoustic image (for example Moushegian & Jeffress, 1959 and Whitworth and Jeffress, 1961). The actual magnitude of the trading ratio obtained in these experiments has been subject to wide variation, depending, inter alia, on the nature of the acoustic signal employed, and the level at which it was presented. Large subject differences were also noted.

A brief historical survey is to be found in Deatherage and Hirsh (1959).

The actual mechanism whereby this trading of time for intensity is accomplished has been variously ascribed to neural latencies, travelling-wave delays in the cochlea or a complex combination of the two. Such a mechanism may be necessary, but would seemingly introduce a lateral shift of the judgement curve as a result of IAD, i.e., along the ITD axis, rather than the largely vertical shift observed in the experiments of section 4.2. In fact, it is hardly feasible to differentiate these two types of shift by centering experiments. The method by which the position of the 'target' image is matched by that of a 'pointer' image, presumes some important knowledge of the characteristics of the 'pointer'. Such information is not yet available.

4.6.0 Conclusions

The experimental method whereby a listener is required to judge the apparent position of a binaurally created sound image has been applied to a number of experimental situations. Results of these experiments have proved to be consistent and reproducible with a number of listeners.

When the technique is applied to certain familiar experiments the results have shown to be rather more instructive than those obtained by other methods. In fact, with a few simple, justified assumptions about image movement with ITD and IAD, and about the process by which positional judgements are elicited, the results of seemingly complex experimental situations may be closely predicted. Certain features of the precedence effect, for example, may be described on the basis of positional judgements of a multiple-image situation.

While the arguments presented here provide a basis for describing certain experimental results, there are still questions to be answered. These relate mainly to the auditory mechanisms whereby these multiple images are established, and further elucidation of the judgement process. The following chapters will describe work aimed at clarifying these issues.

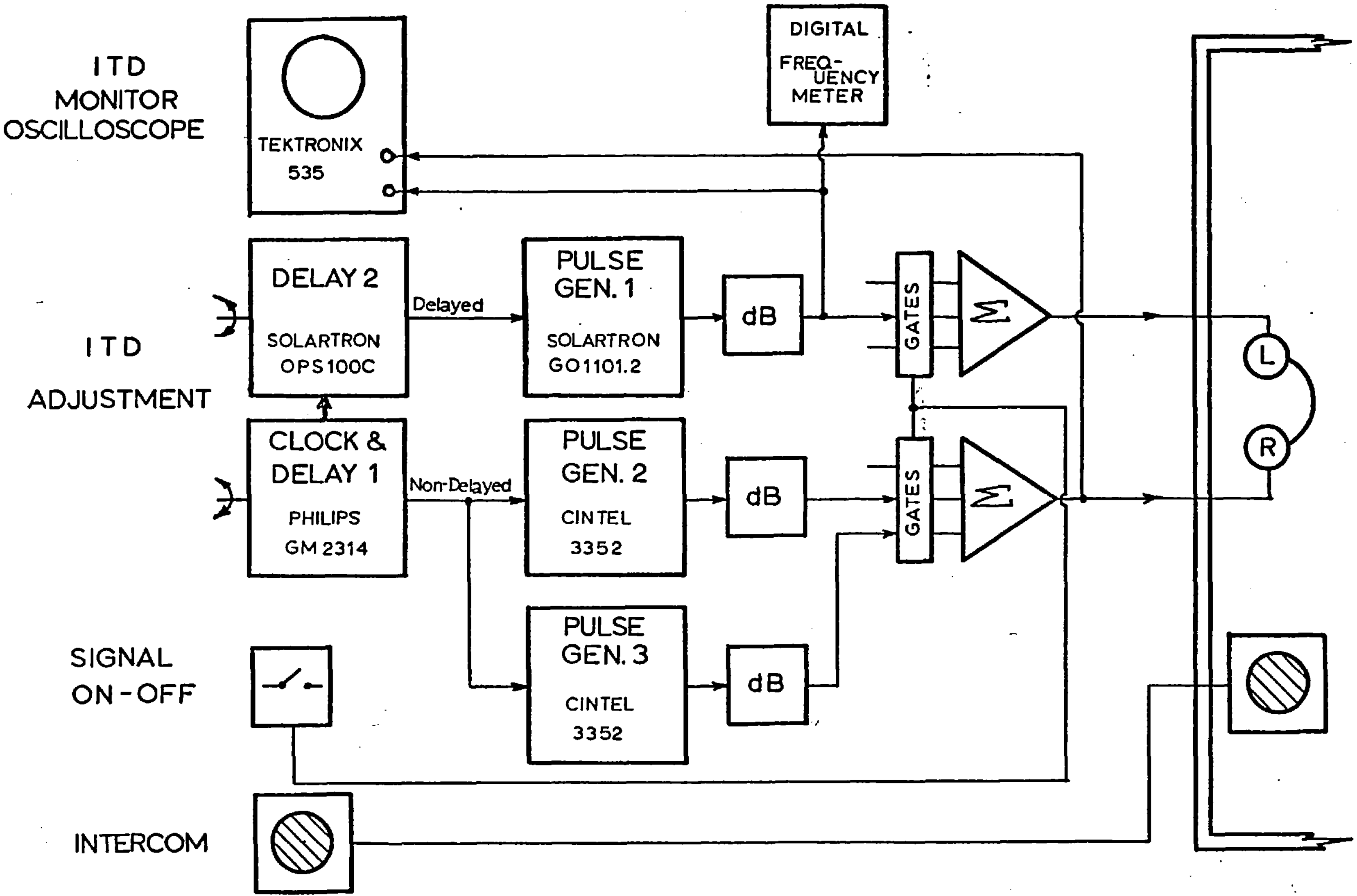

ITD
MONITOR
OSCILLOSCOPE
TEKTRONIX
535
DIGITAL
FREQ-
UENCY
METER
ITD
ADJUSTMENT
DELAY 2
SOLARTRON
OPS100C
Delayed
PULSE
GEN. 1
SOLARTRON
GO1101.2
dB
GATES
Σ
L
R
CLOCK &
DELAY 1
PHILIPS
GM 2314
Non-Delayed
PULSE
GEN. 2
CINTEL
3352
dB
GATES
Σ
SIGNAL
ON-OFF
PULSE
GEN. 3
CINTEL
3352
dB
INTERCOM

FIGURE 4.1

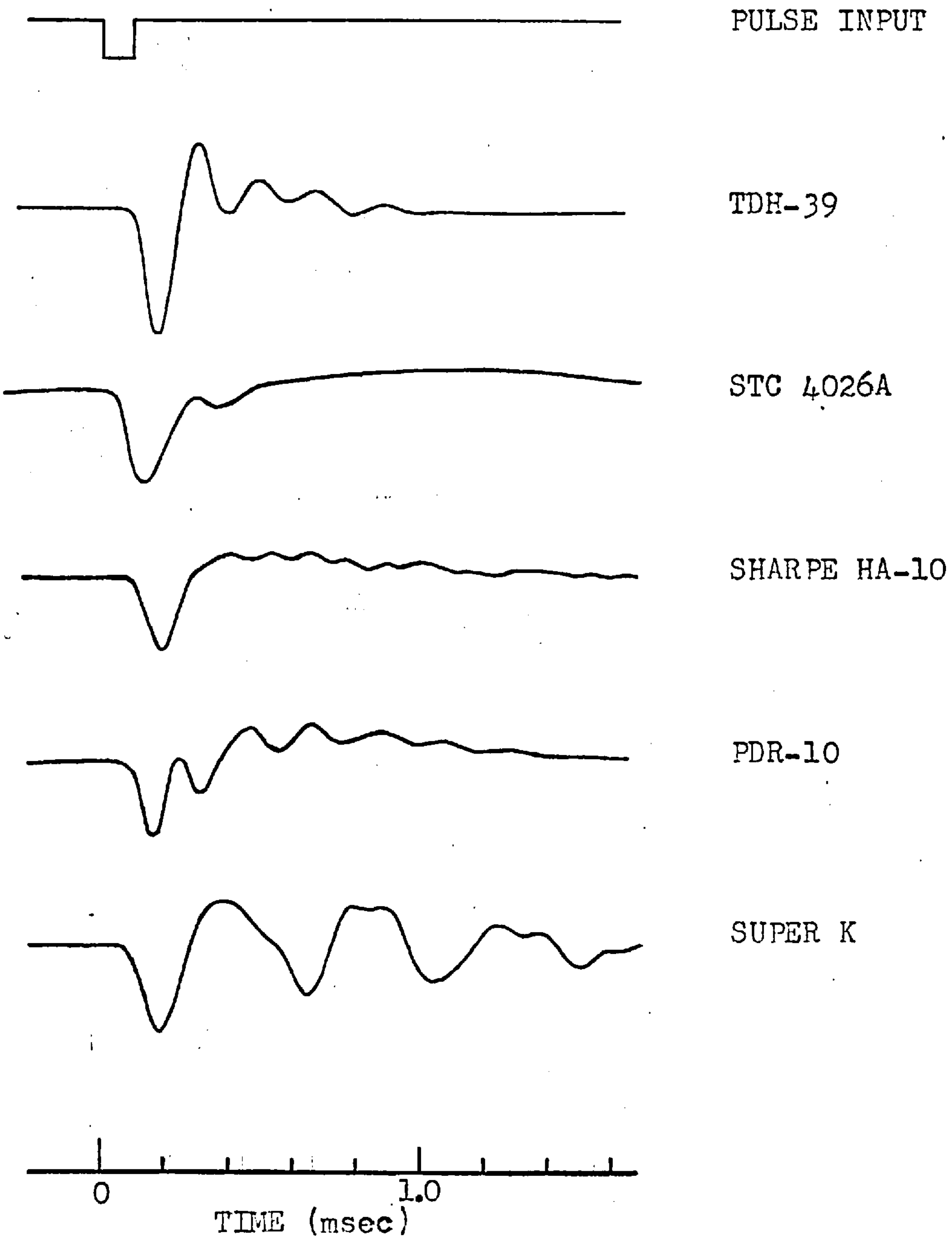
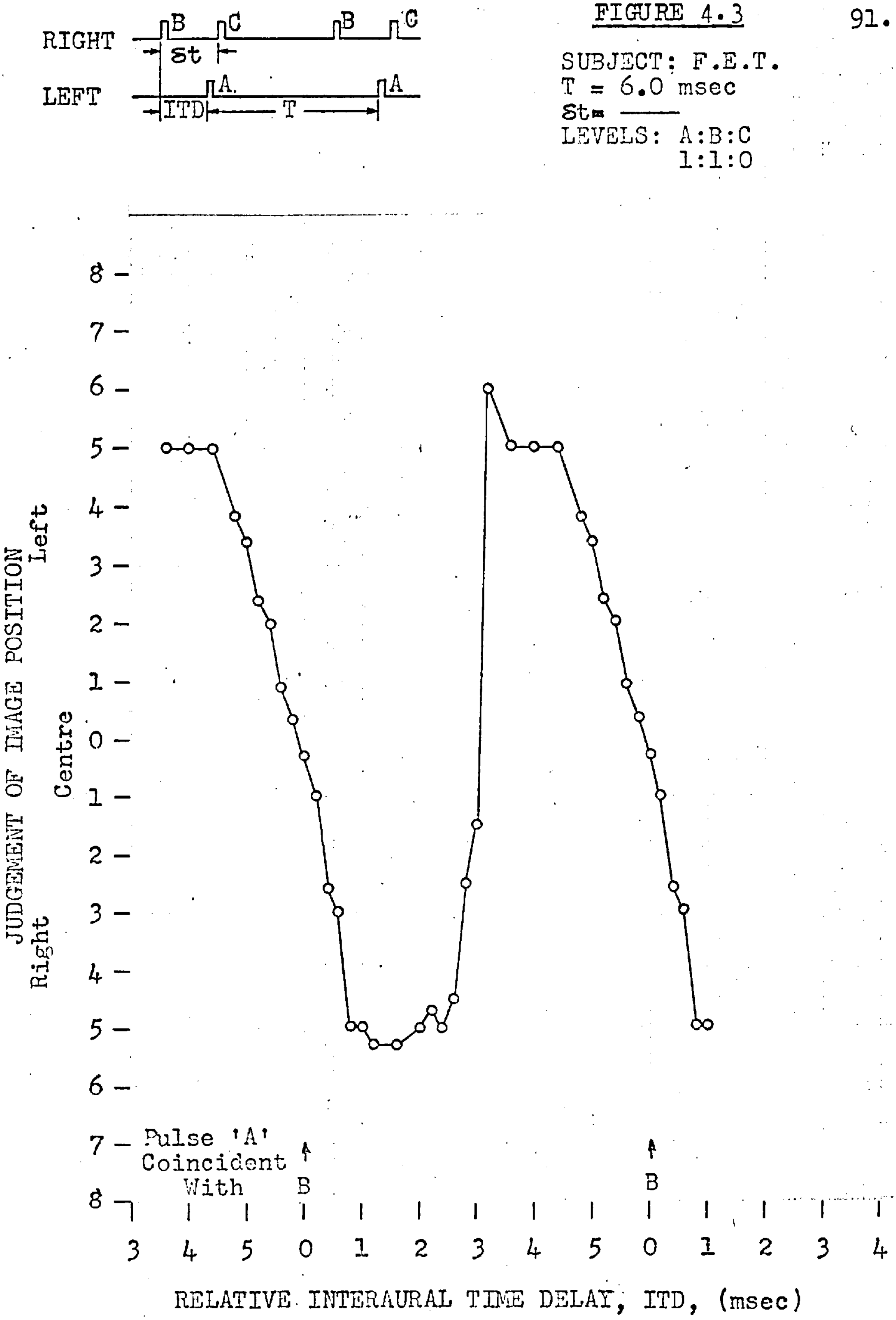


FIGURE 4.2 : Waveforms of acoustic transients produced by various earphones. Measurements were made using an artificial ear.

FIGURE 4.3



SUBJECT: F.E.T.
T = 6.0 msec
δt= ——
LEVELS: A:B:C
1:1:0

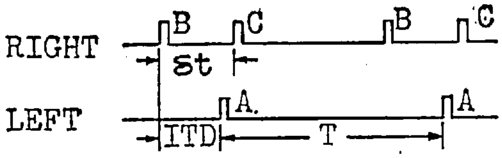
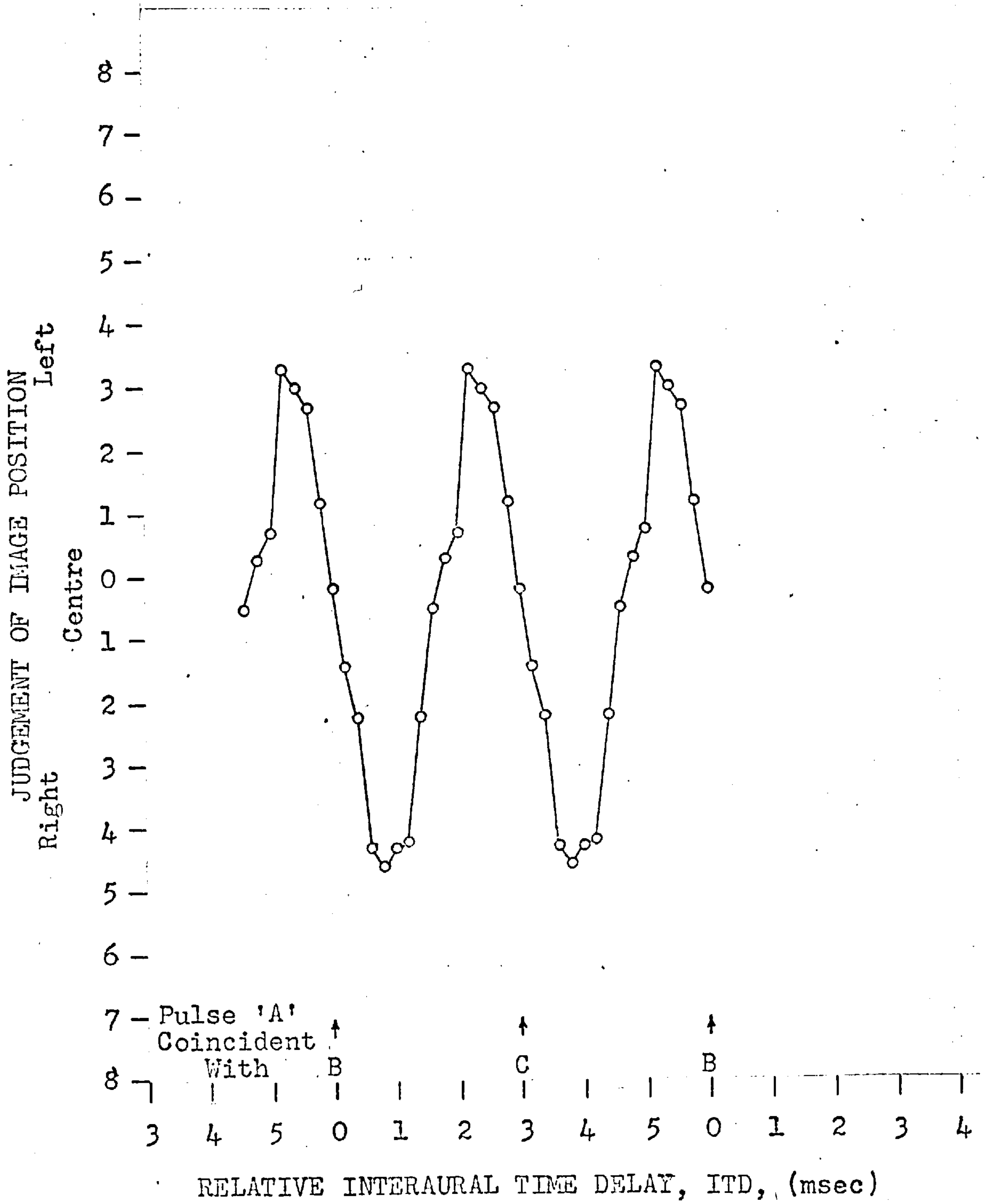


FIGURE 4.4

SUBJECT: F.E.T.
T = 6.0 msec
δt = 3.0 msec
LEVELS: A:B:C
1:1:1



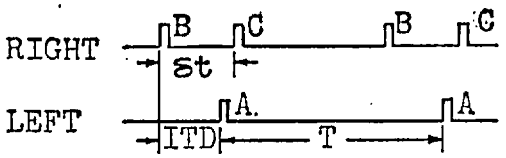


FIGURE 4.5

SUBJECT: F.E.T.
T = 6.0 msec
δt= 2.7 msec
LEVELS: A:B:C:
1:1:1

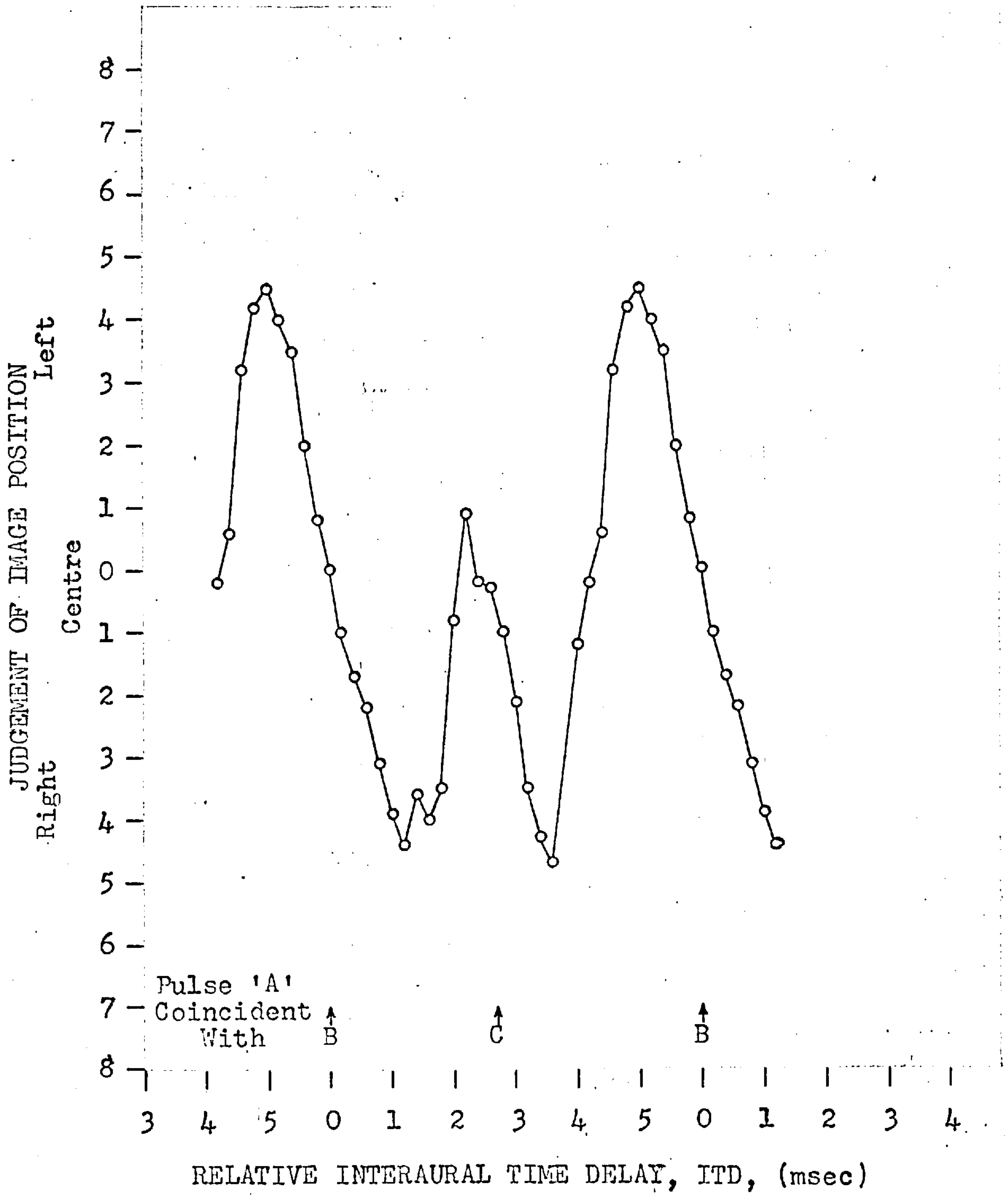
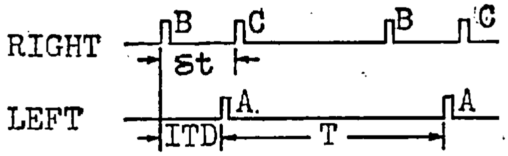
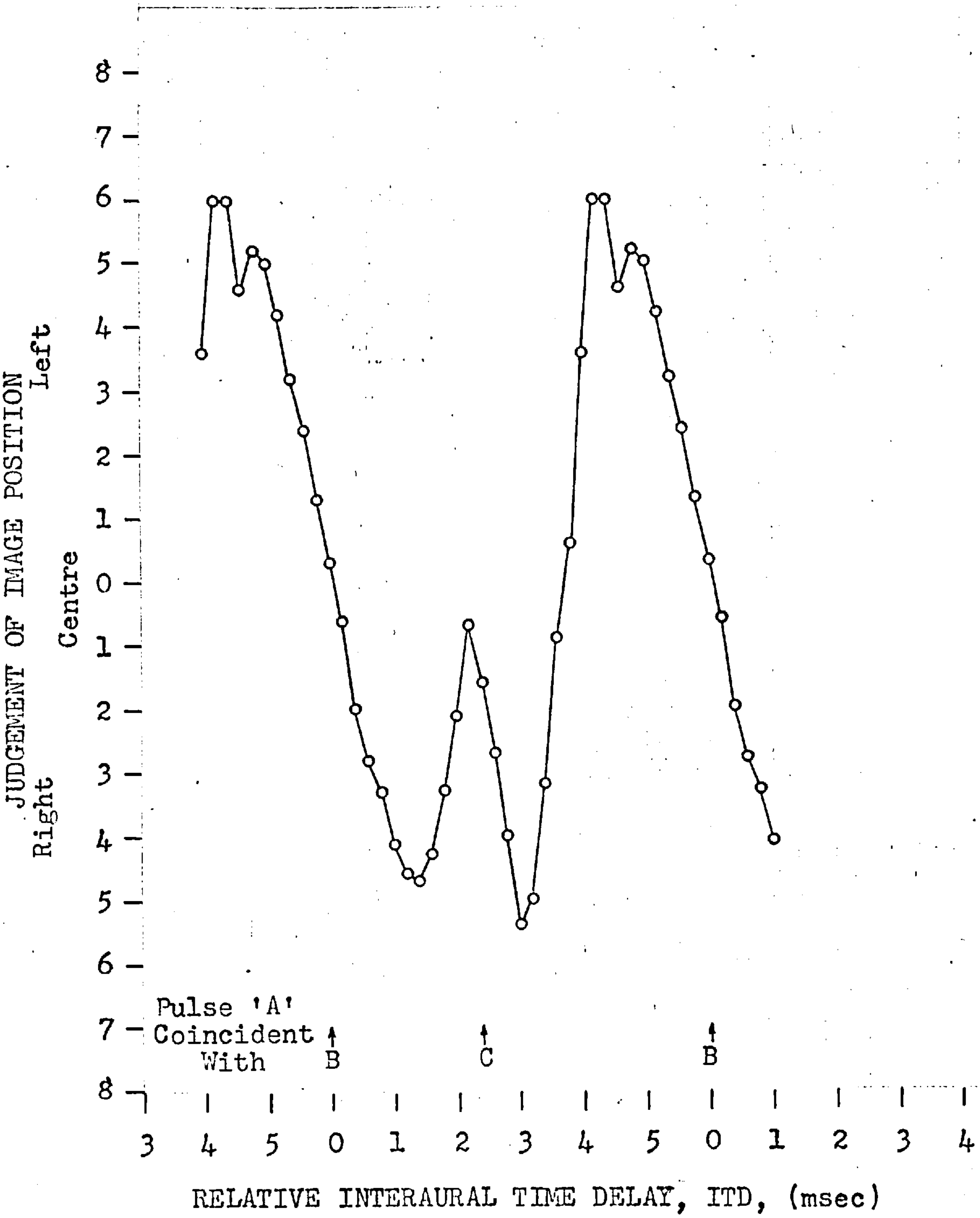


FIGURE 4.6

SUBJECT: F.E.T.
T = 6.0 msec
δt= 2.4 msec
LEVELS: A:B:C
1:1:1





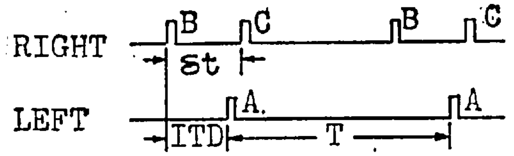
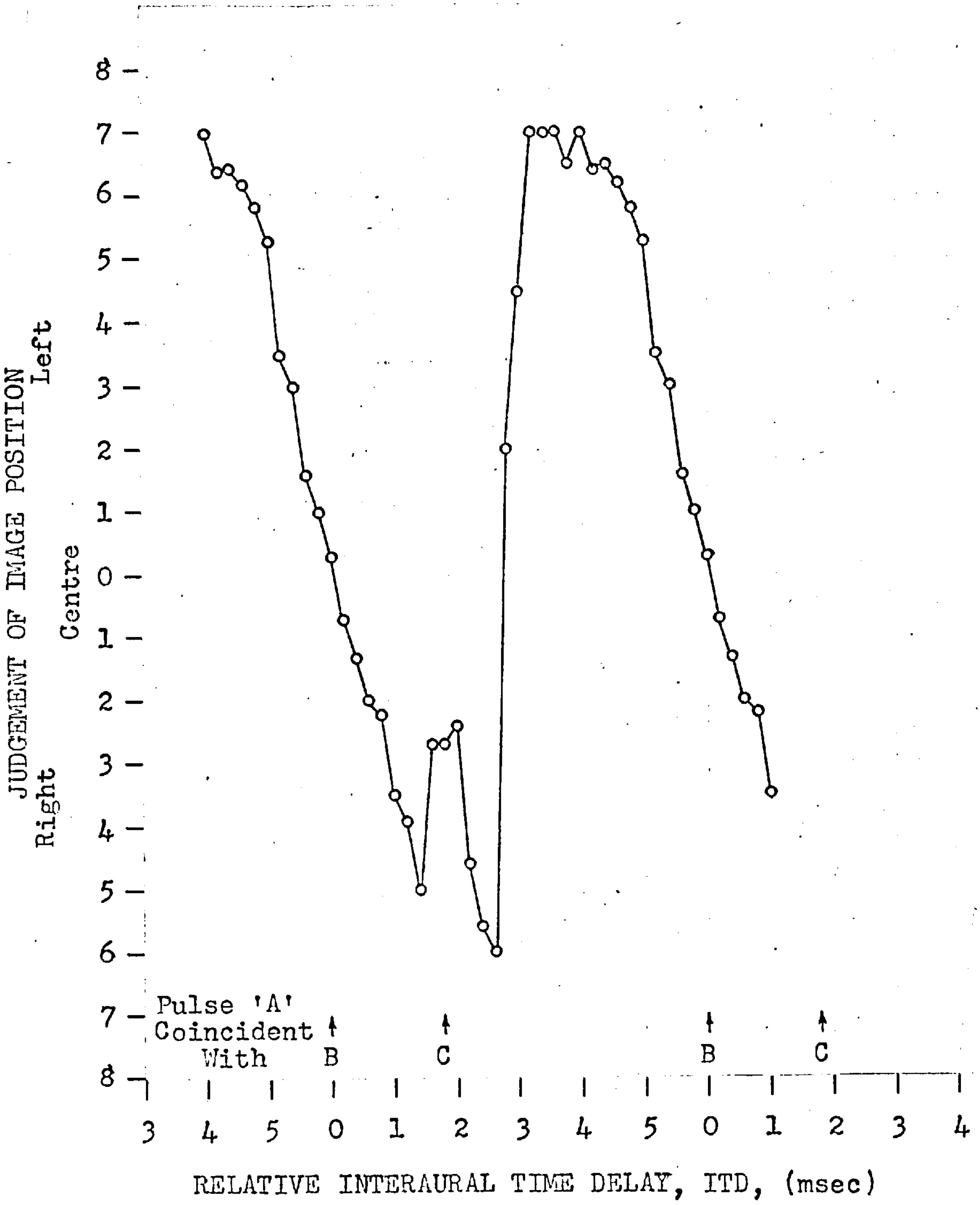


FIGURE 4.7

SUBJECT: F.E.T.
T = 6.0 msec
δt= 1.8 msec
LEVELS: A:B:C
1:1:1



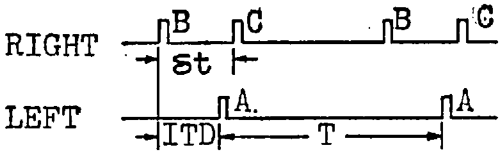


FIGURE 4.8

SUBJECT: F.E.T.
T = 6.0 msec
δt = 0.6 msec
LEVELS: A:B:C
1:1:1

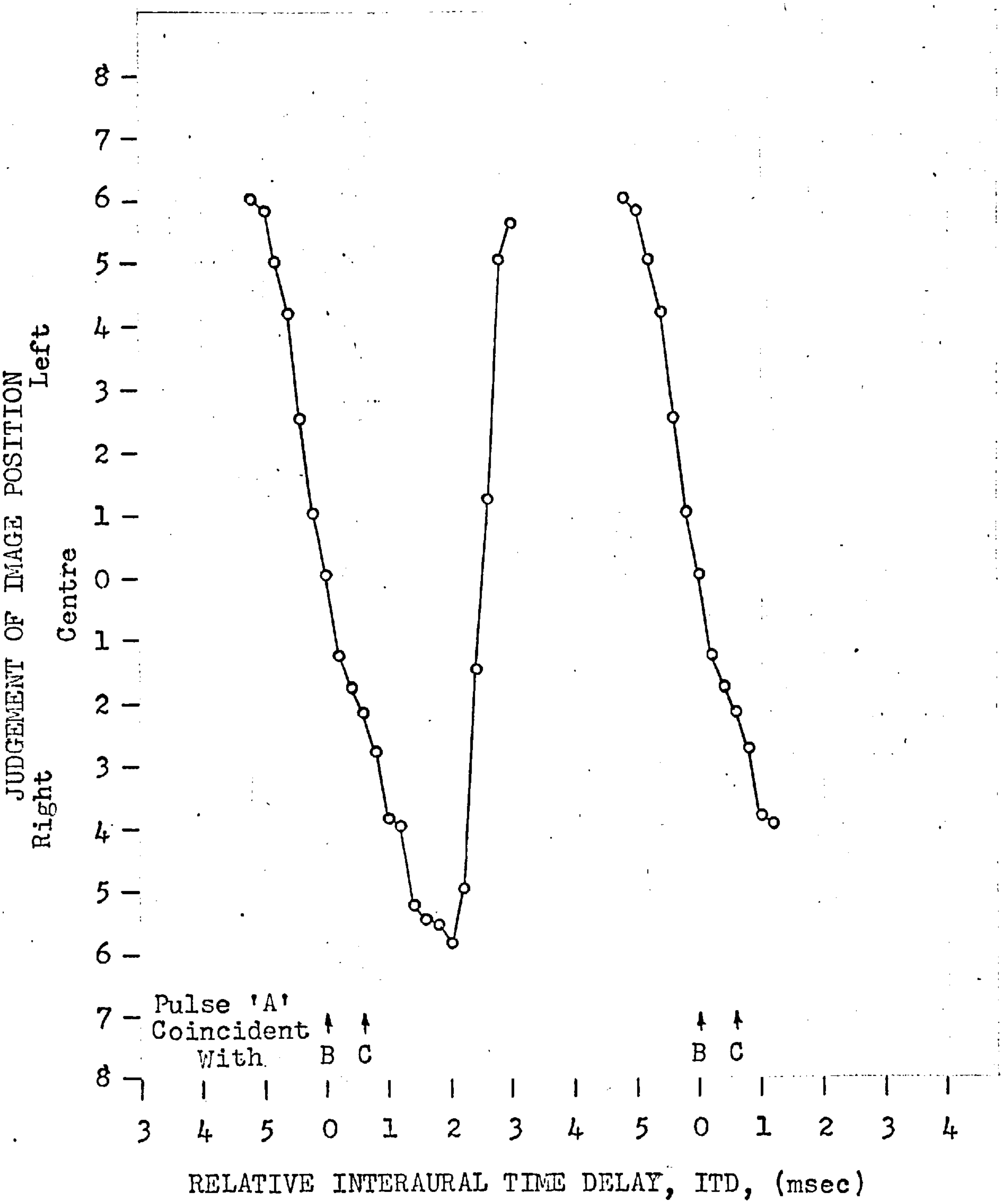
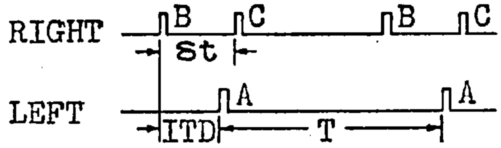


FIGURE 4.9



SUBJECT: F.E.T.
T = 10.0 msec
δt= 5.0 msec
LEVELS: A:B:C
1:1:1

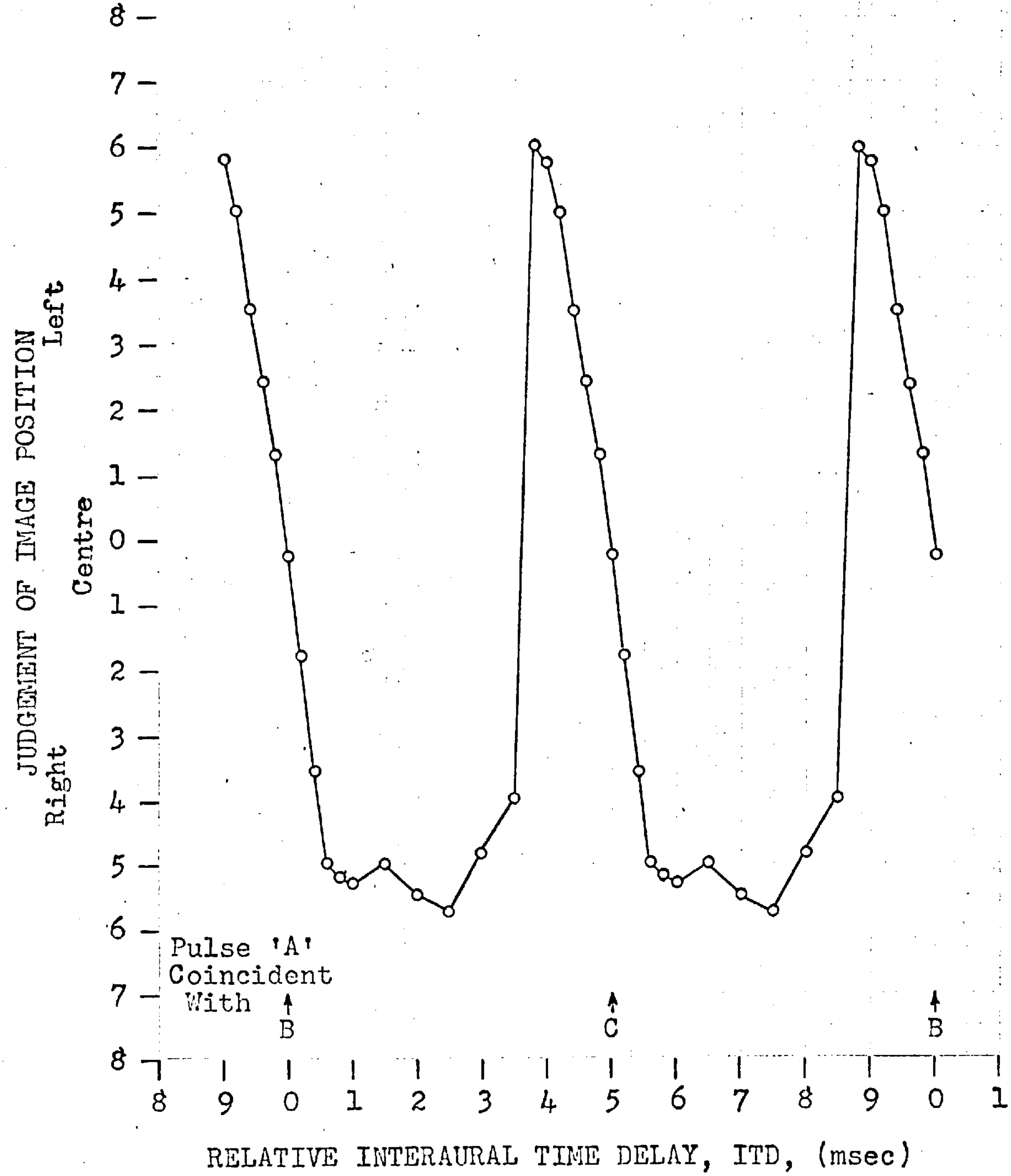


FIGURE 4.10

SUBJECT: F.E.T.
T = 10.0 msec
δt= 3.0 msec
LEVELS: A:B:C
1:1:1

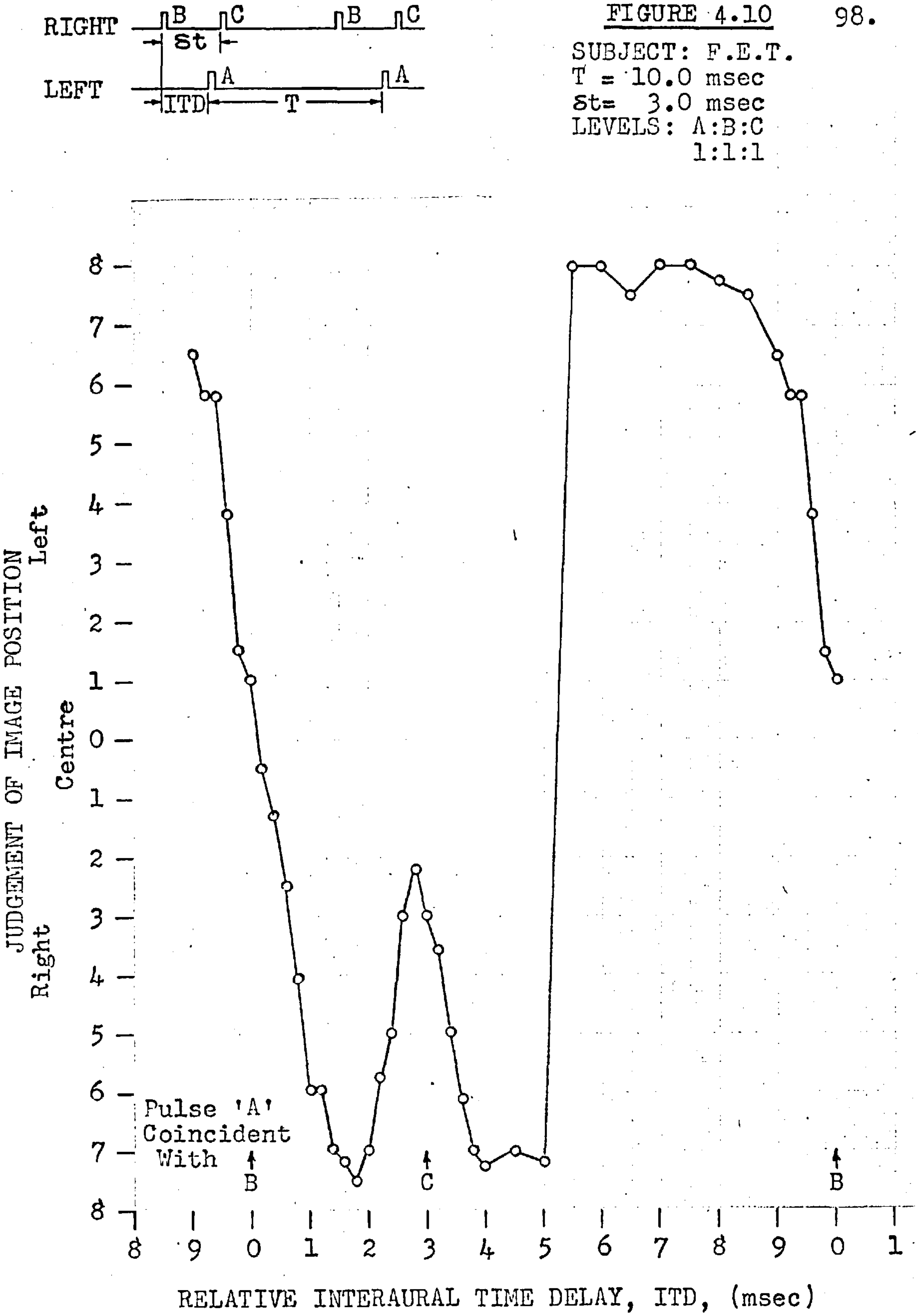
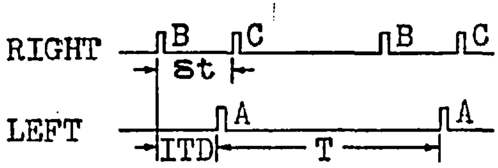
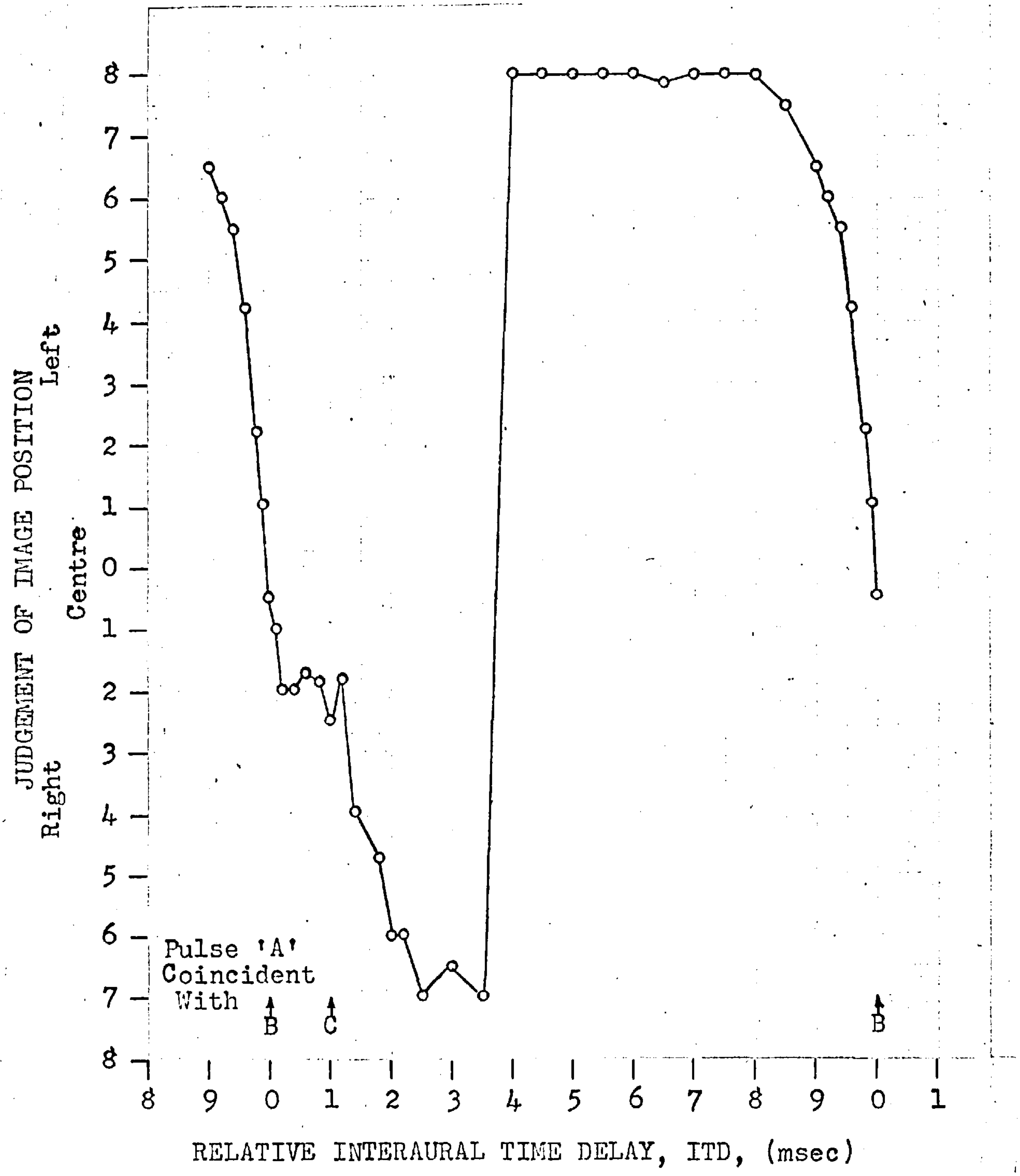


FIGURE 4.11

SUBJECT: F.E.T.
T = 10.0 msec
δt= 1.0 msec
LEVELS: A:B:C
1:1:1





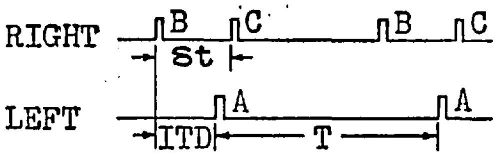


FIGURE 4.12

SUBJECT: J.P.B.
T = 10.0 msec
δt= 5.0 msec
LEVELS: A:B:C
1:1:1

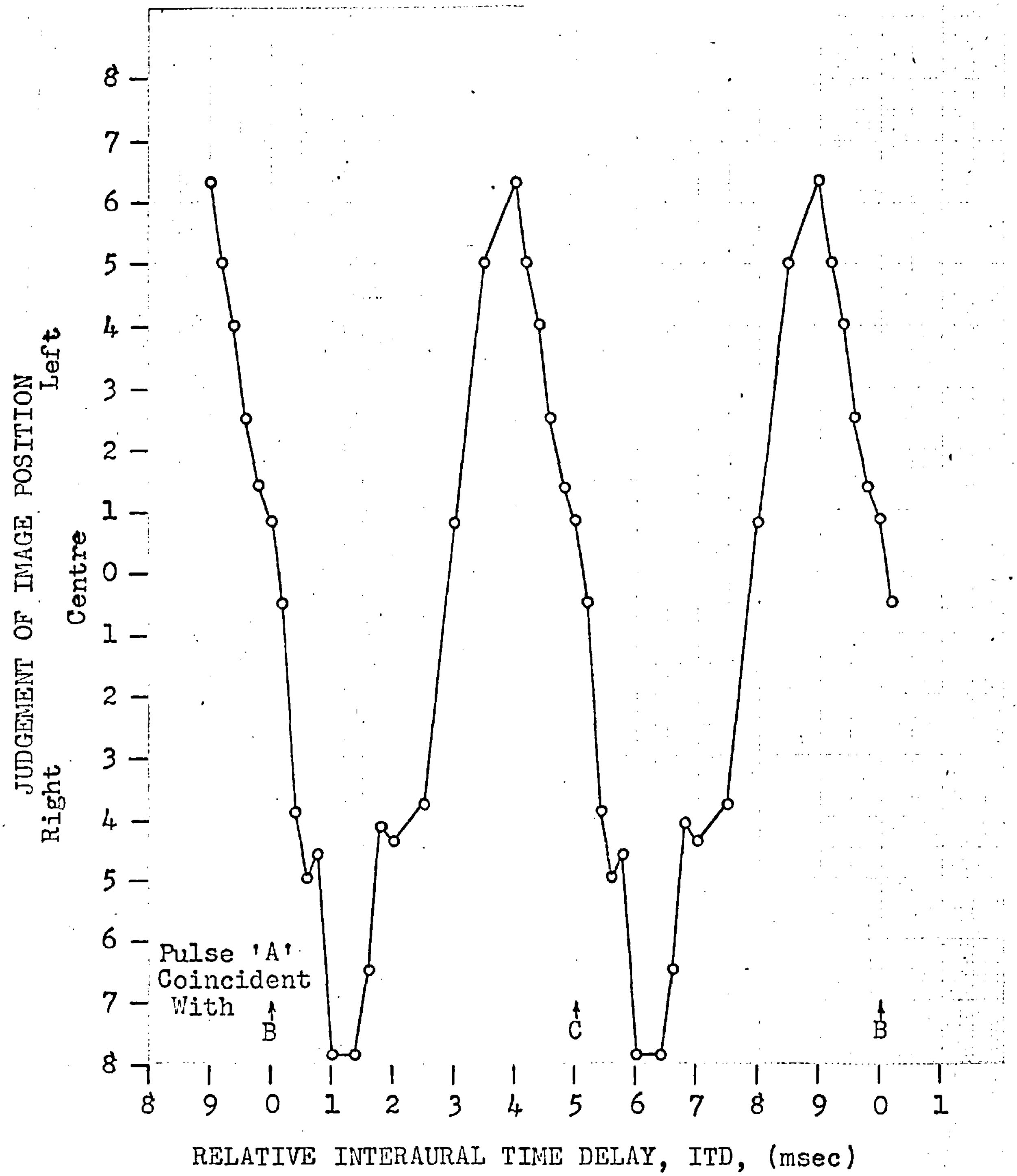


FIGURE 4.13

SUBJECT: J.P.B.
T = 10.0 msec
δt= 3.0 msec
LEVELS: A:B:C
1:1:1

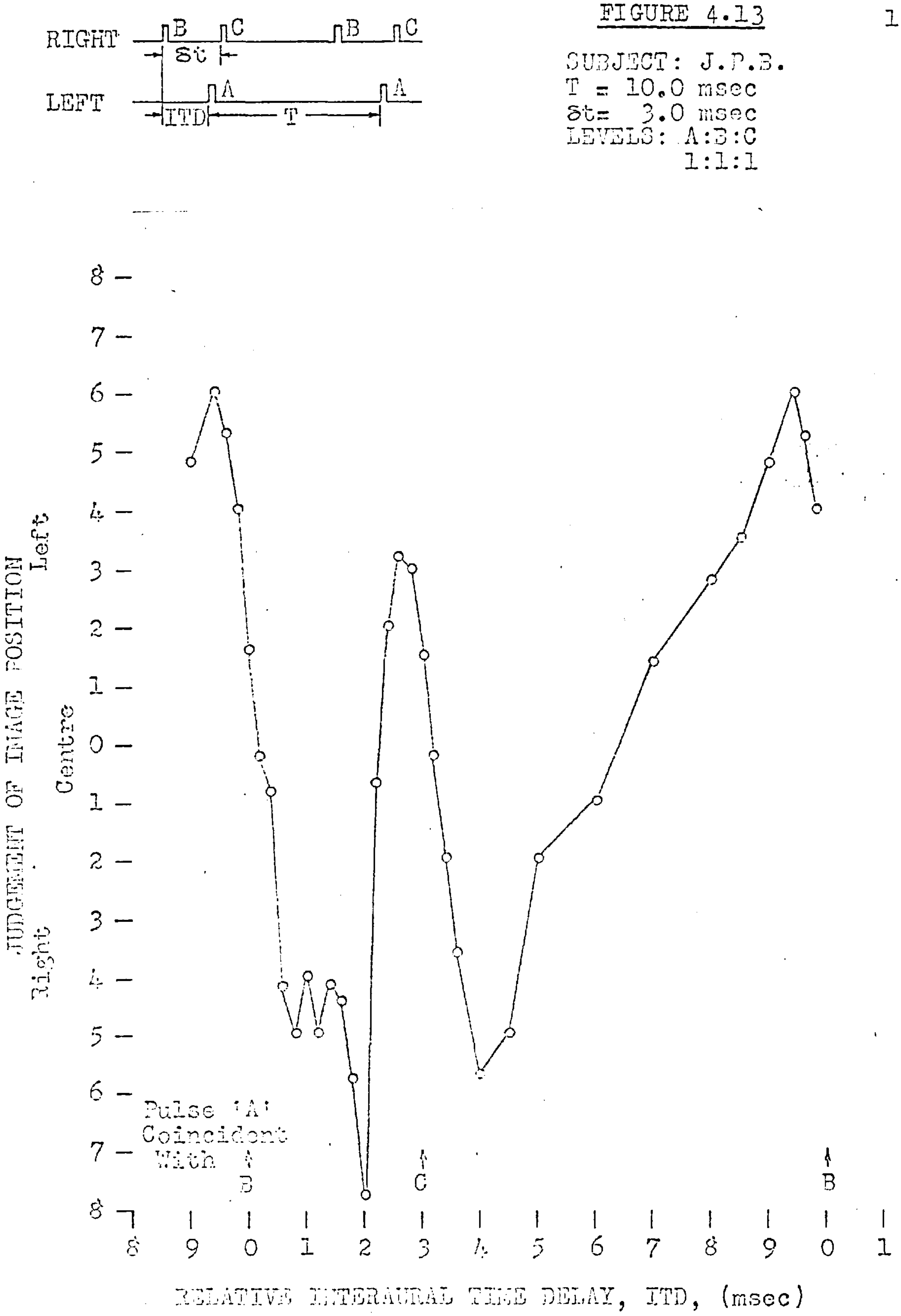


FIGURE 4.14

SUBJECT: J.P.B.
T = 10.0 msec
δt = 1.0 msec
LEVELS: A:B:C
1:1:1

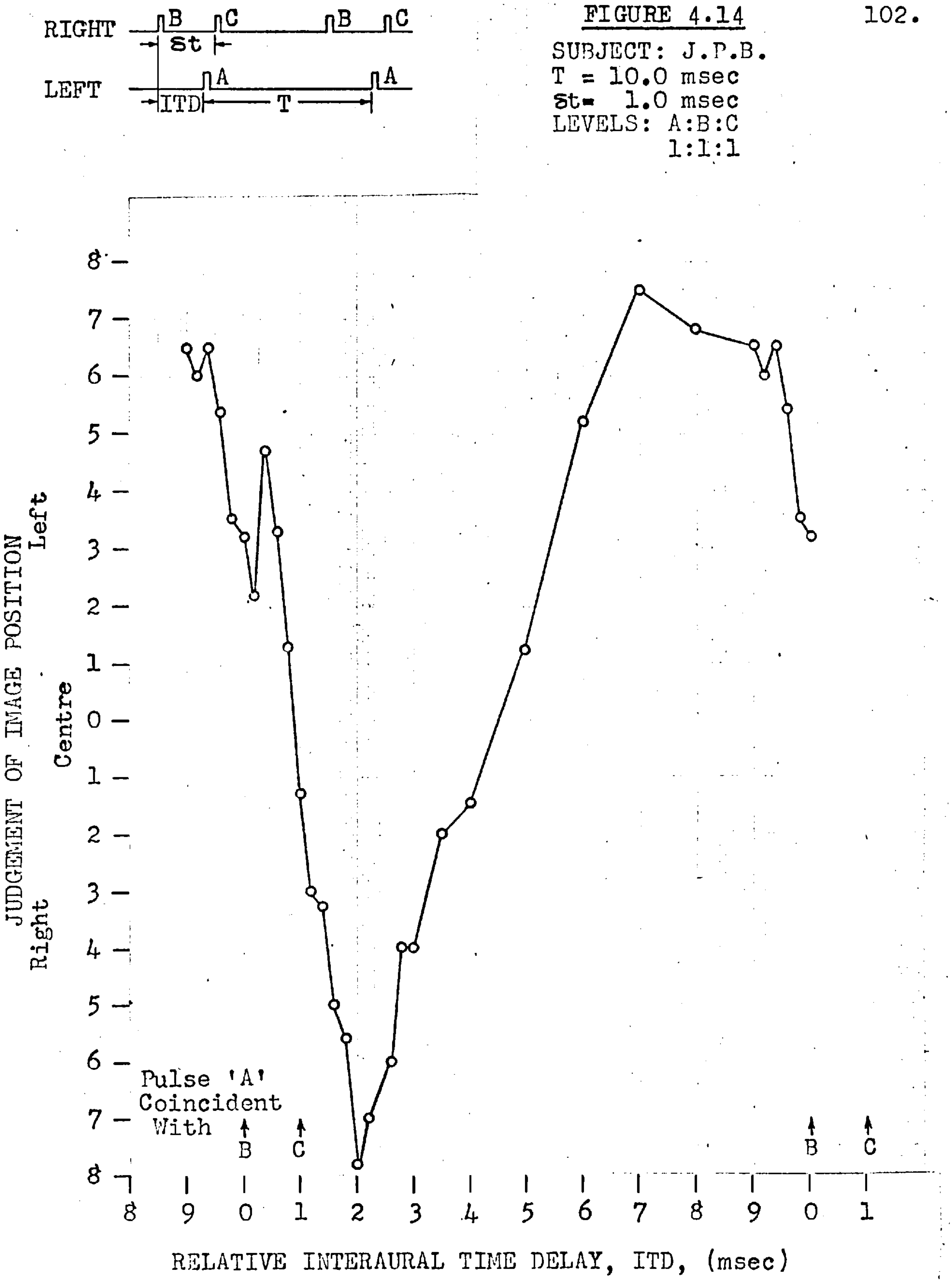
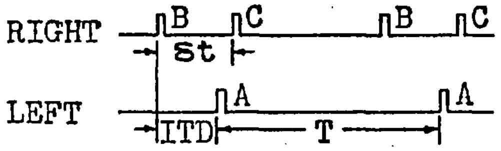


FIGURE 4.15

SUBJECT: R.E.T.
T = 10.0 msec
δt= 5.0 msec
LEVELS: A:B:C
1:1:1



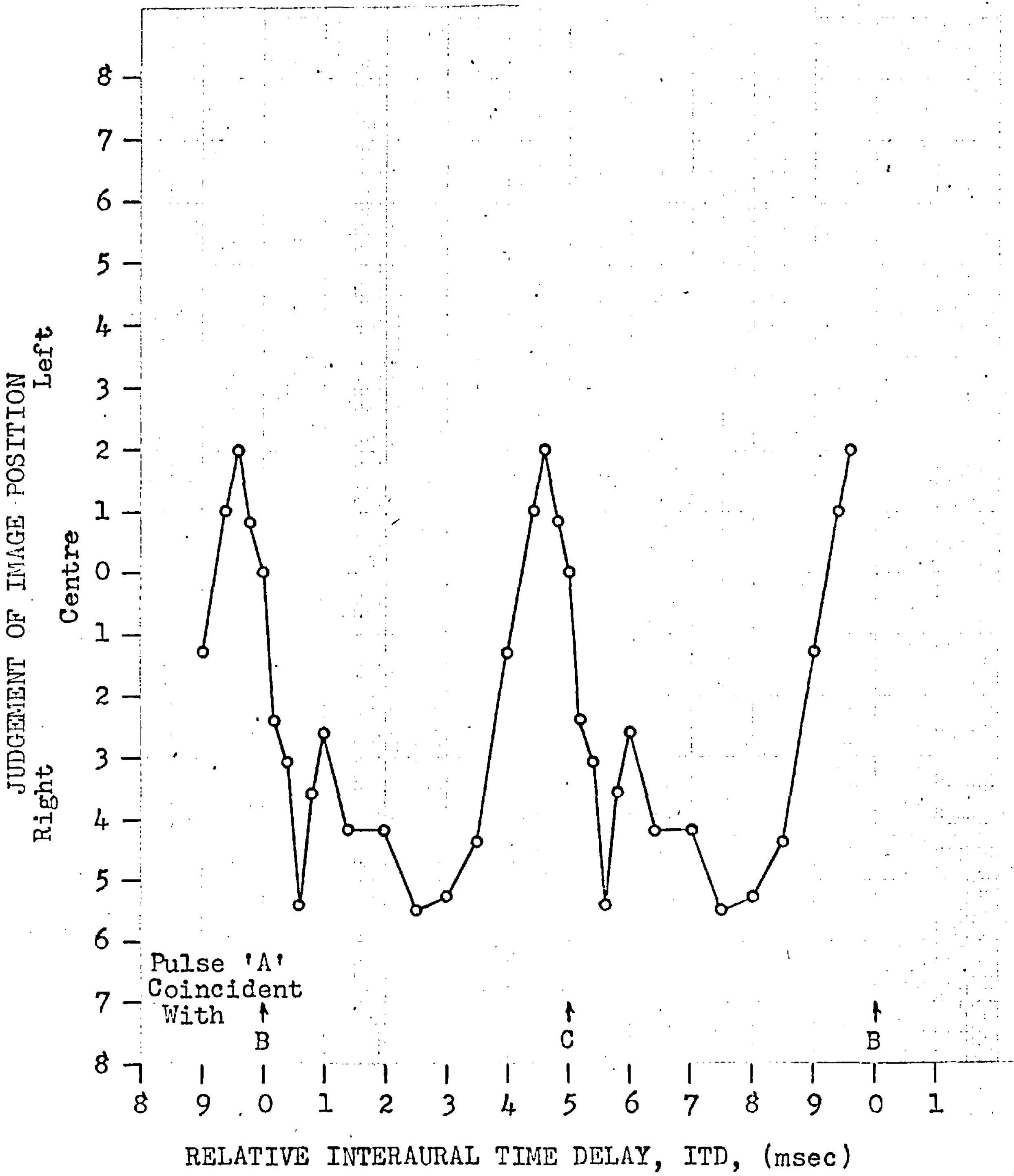
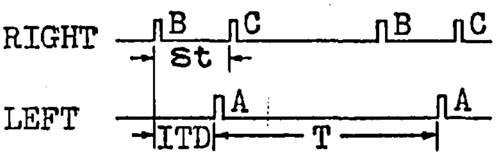
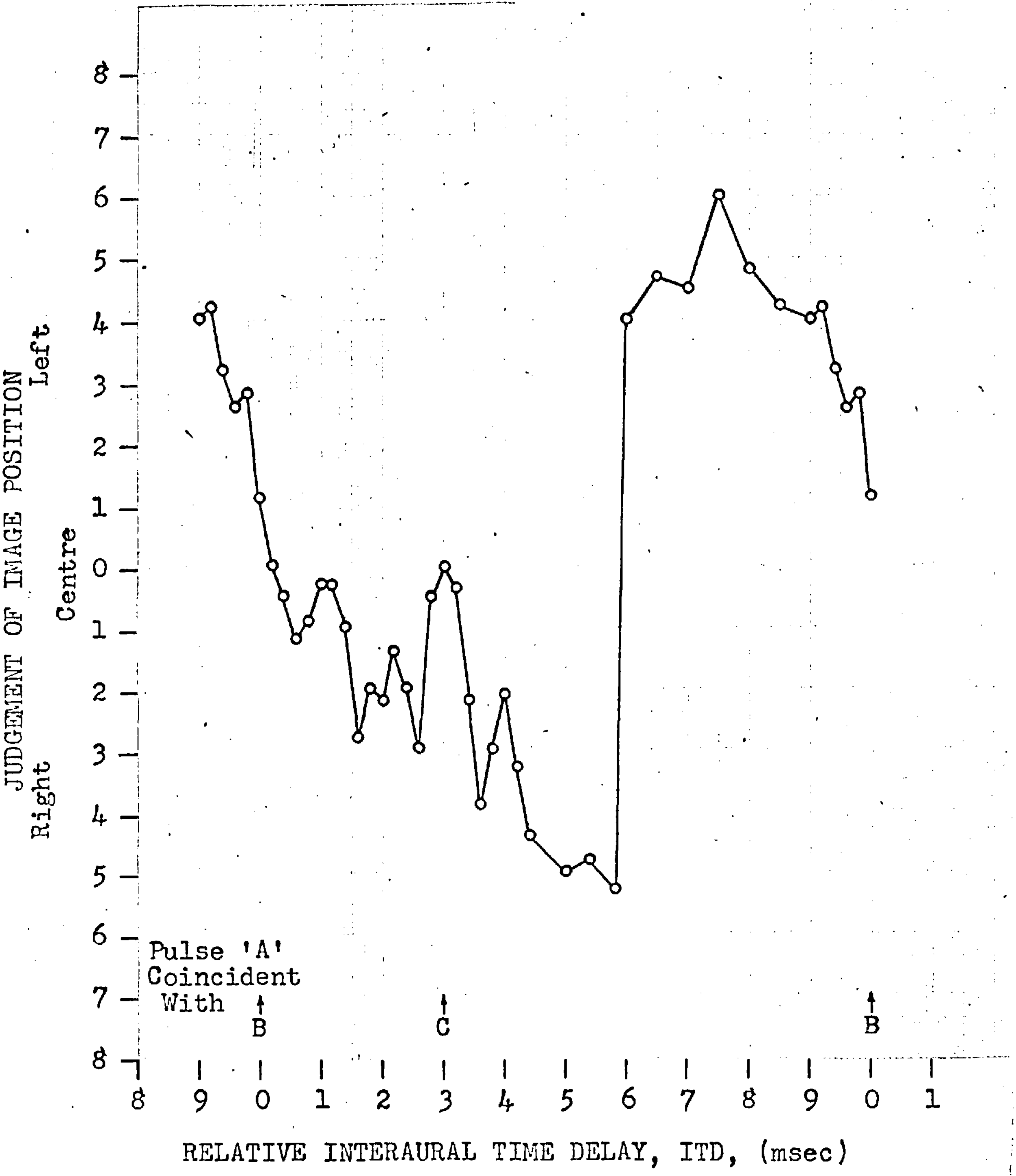


FIGURE 4.16

SUBJECT: R.E.T.
T = 10.0 msec
δt= 3.0 msec
LEVELS: A:B:C
1:1:1





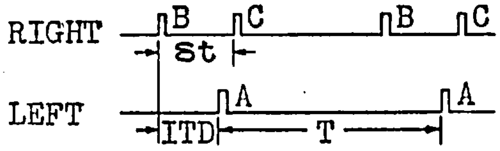


FIGURE 4.17

SUBJECT: R.E.T.
T = 10.0 msec
δt= 1.0 msec
LEVELS: A:B:C
1:1:1

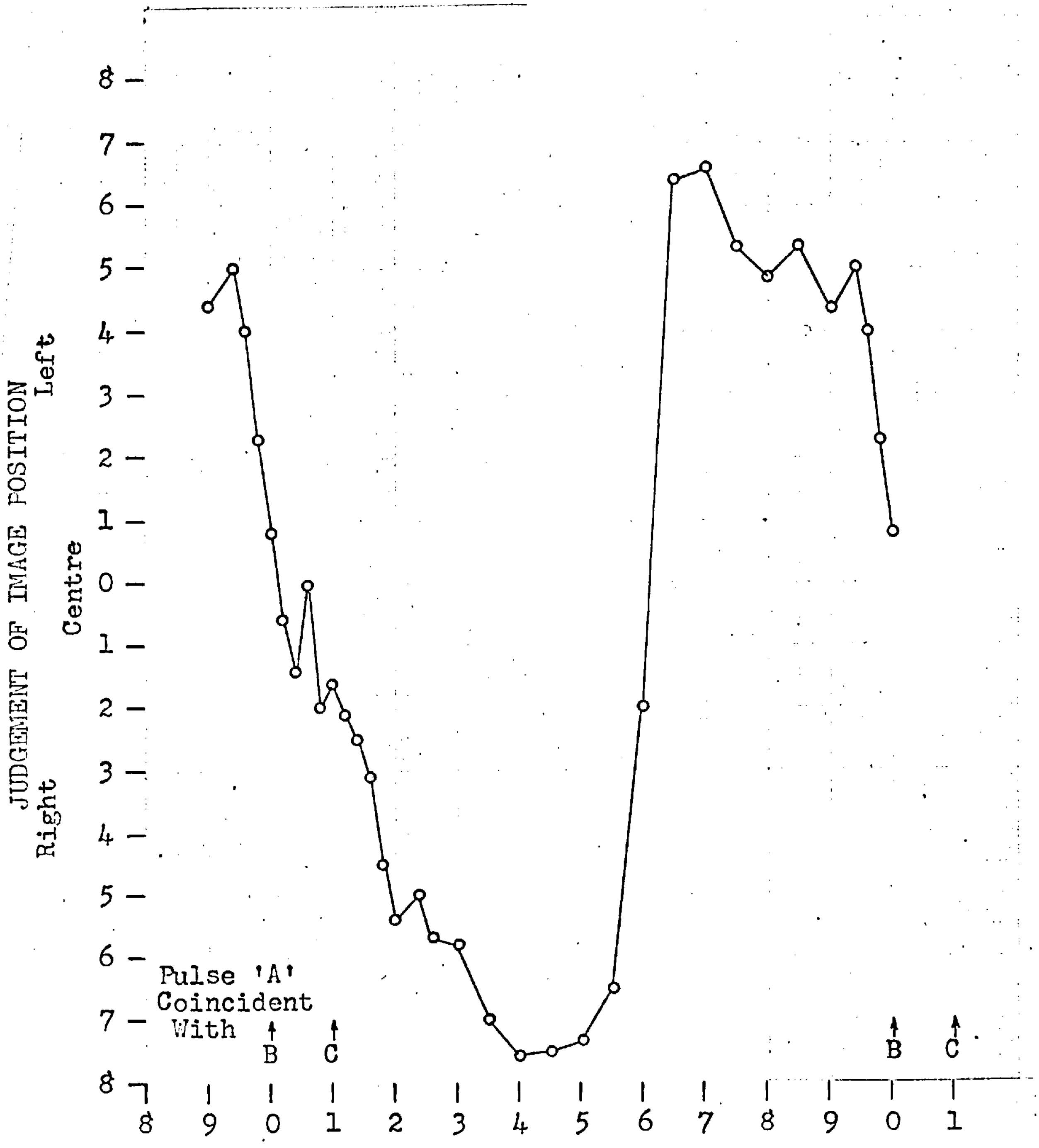


FIGURE 4.18

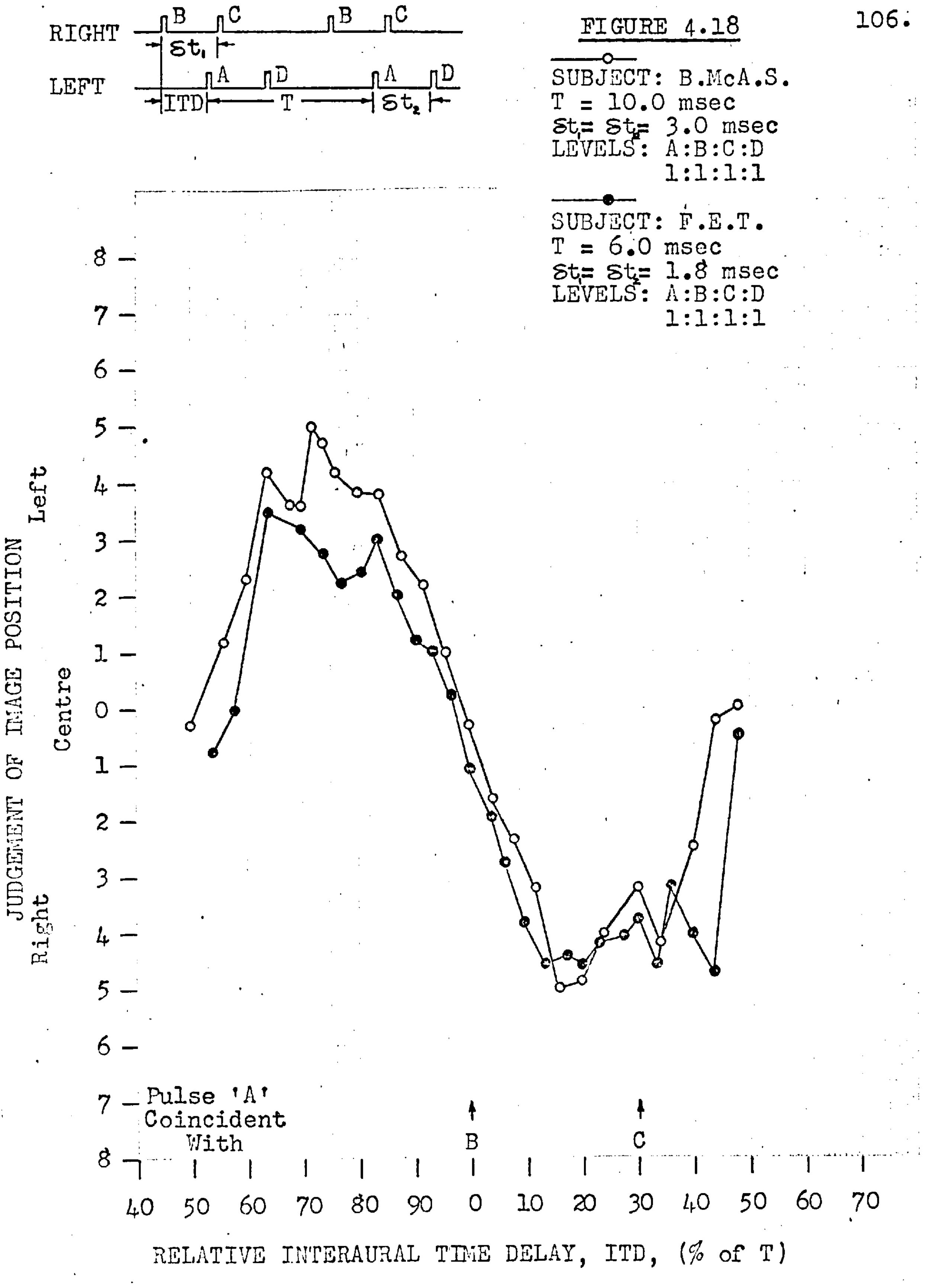

RIGHT
B
C
B
C
δt₁
LEFT
A
D
A
D
ITD
T
δt₂
SUBJECT: B.McA.S.
T = 10.0 msec
δt₁= δt₂= 3.0 msec
LEVELS: A:B:C:D
1:1:1:1
SUBJECT: F.E.T.
T = 6.0 msec
δt₁= δt₂= 1.8 msec
LEVELS: A:B:C:D
1:1:1:1
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
C
40
50
60
70
80
90
0
10
20
30
40
50
60
70
RELATIVE INTERAURAL TIME DELAY, ITD, (% of T)

FIGURE 4.19

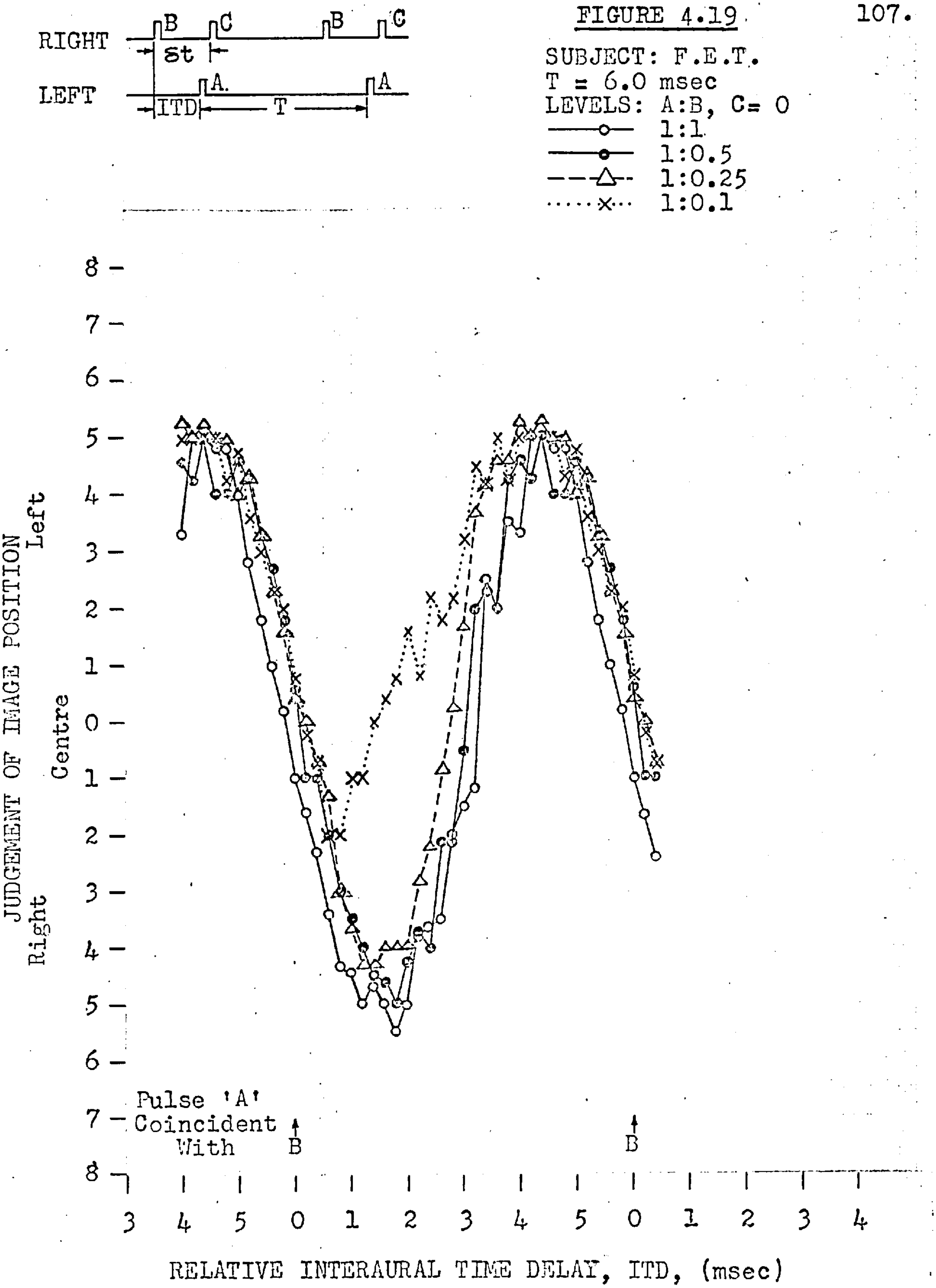

RIGHT
B
C
B
C
δt
LEFT
A.
A
ITD
T
SUBJECT: F.E.T.
T = 6.0 msec
LEVELS: A:B, C= 0
—○— 1:1
—●— 1:0.5
--△-- 1:0.25
···×··· 1:0.1
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
B
3
4
5
0
1
2
3
4
5
0
1
2
3
4
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 4.20

SUBJECT: B.McA.S.
T = 3.0 msec
LEVELS: A:B, C= 0

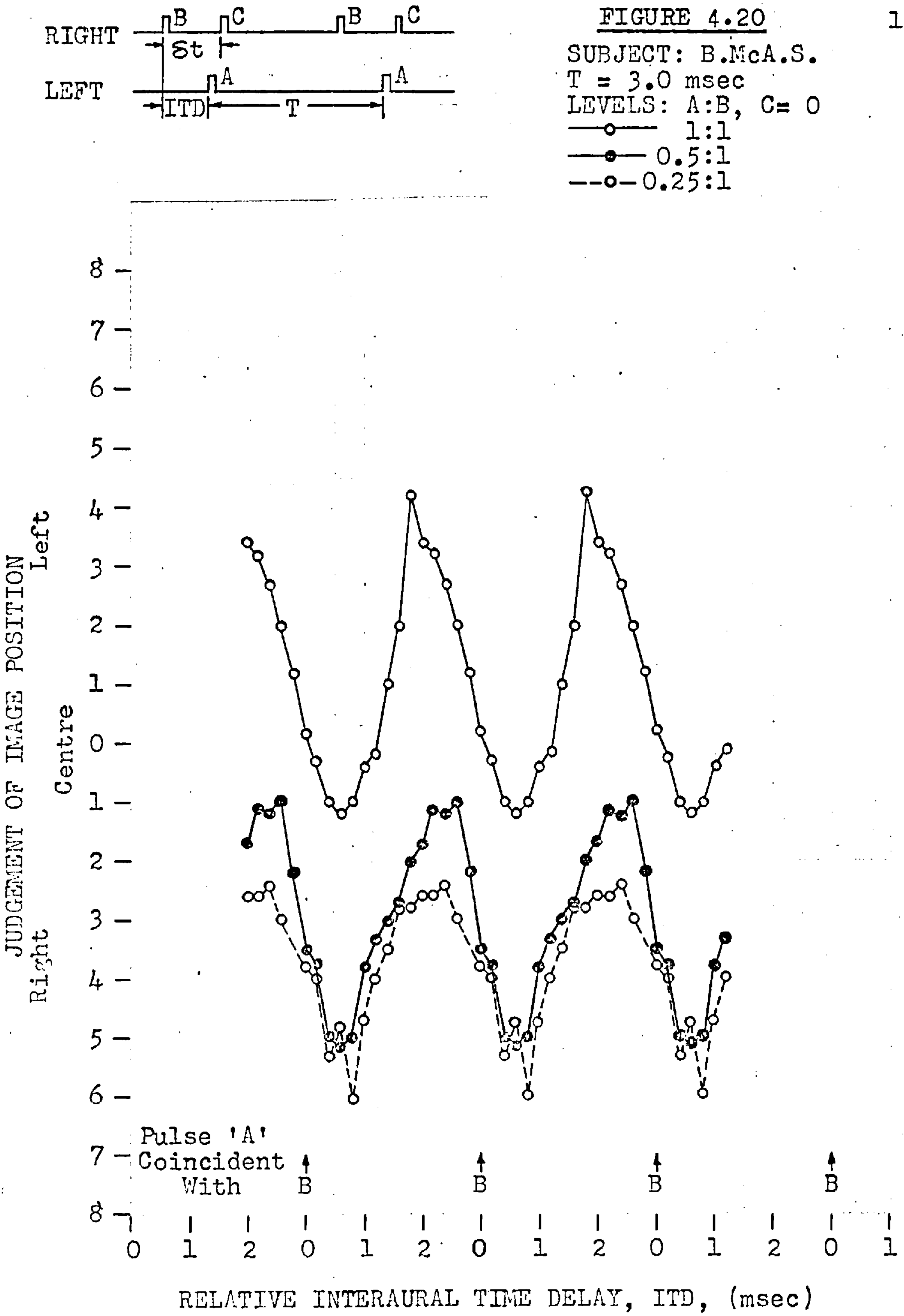
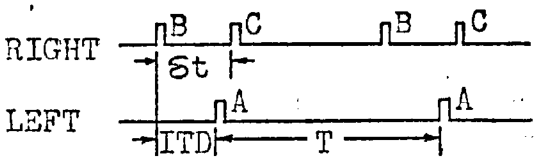


FIGURE 4.21

SUBJECT: F.E.T
T = 3.0 msec
LEVELS: A:B, C= 0
⋯○⋯ 1:1
—●— 0.5:1
—○—0.17:1



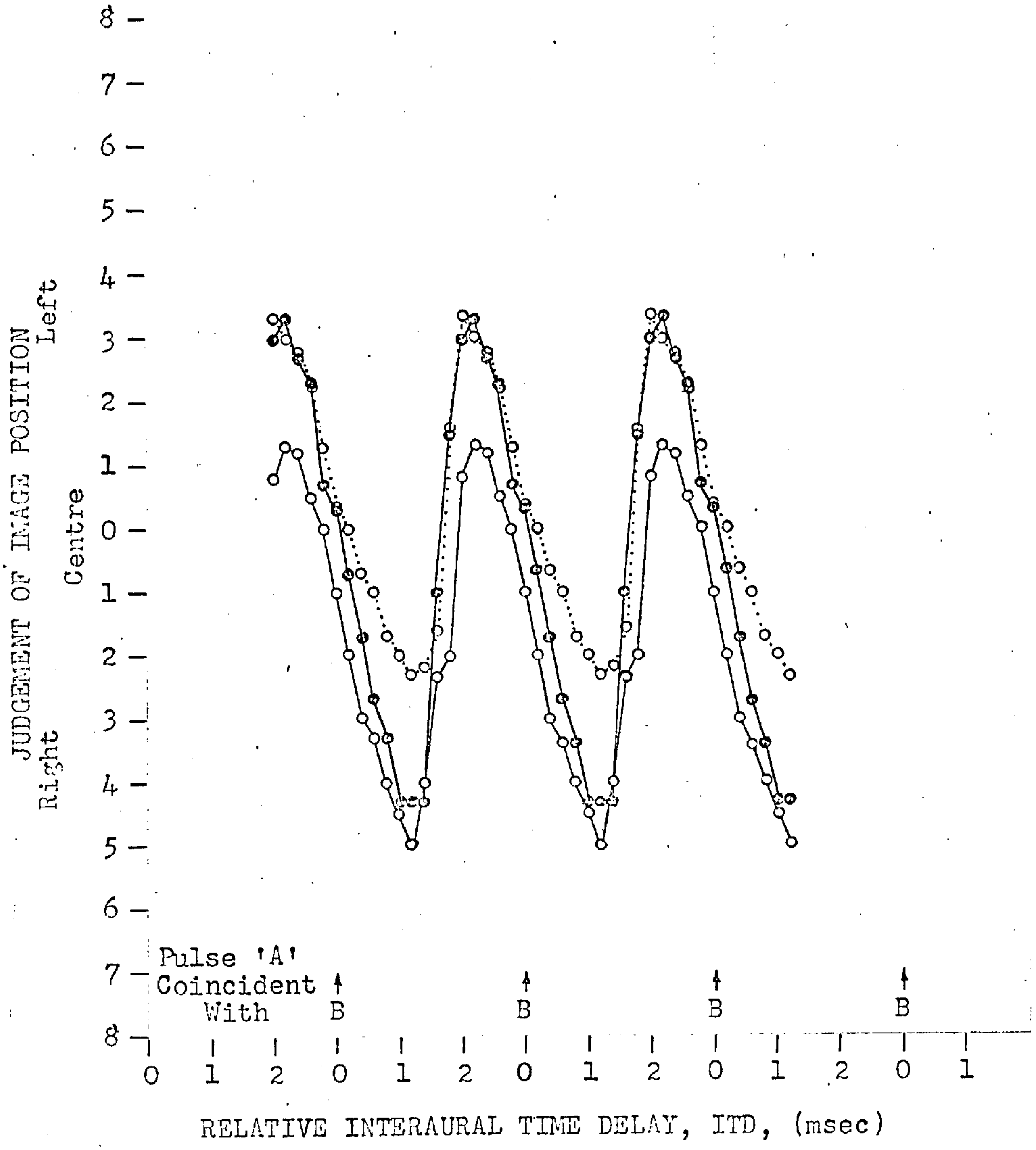
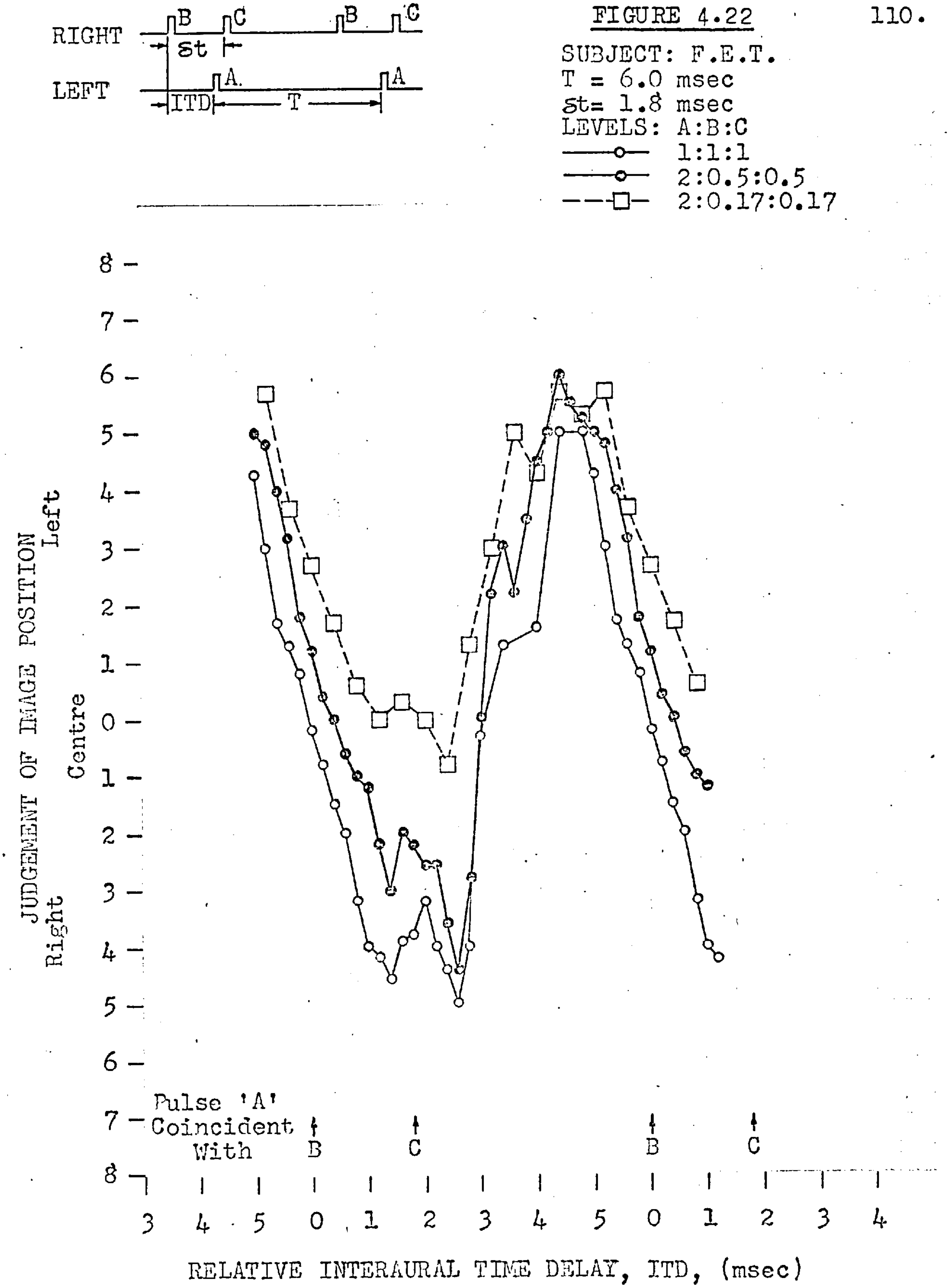


FIGURE 4.22



SUBJECT: F.E.T.
T = 6.0 msec
δt= 1.8 msec
LEVELS: A:B:C
—○— 1:1:1
—●— 2:0.5:0.5
---□- 2:0.17:0.17

FIGURE 4.23

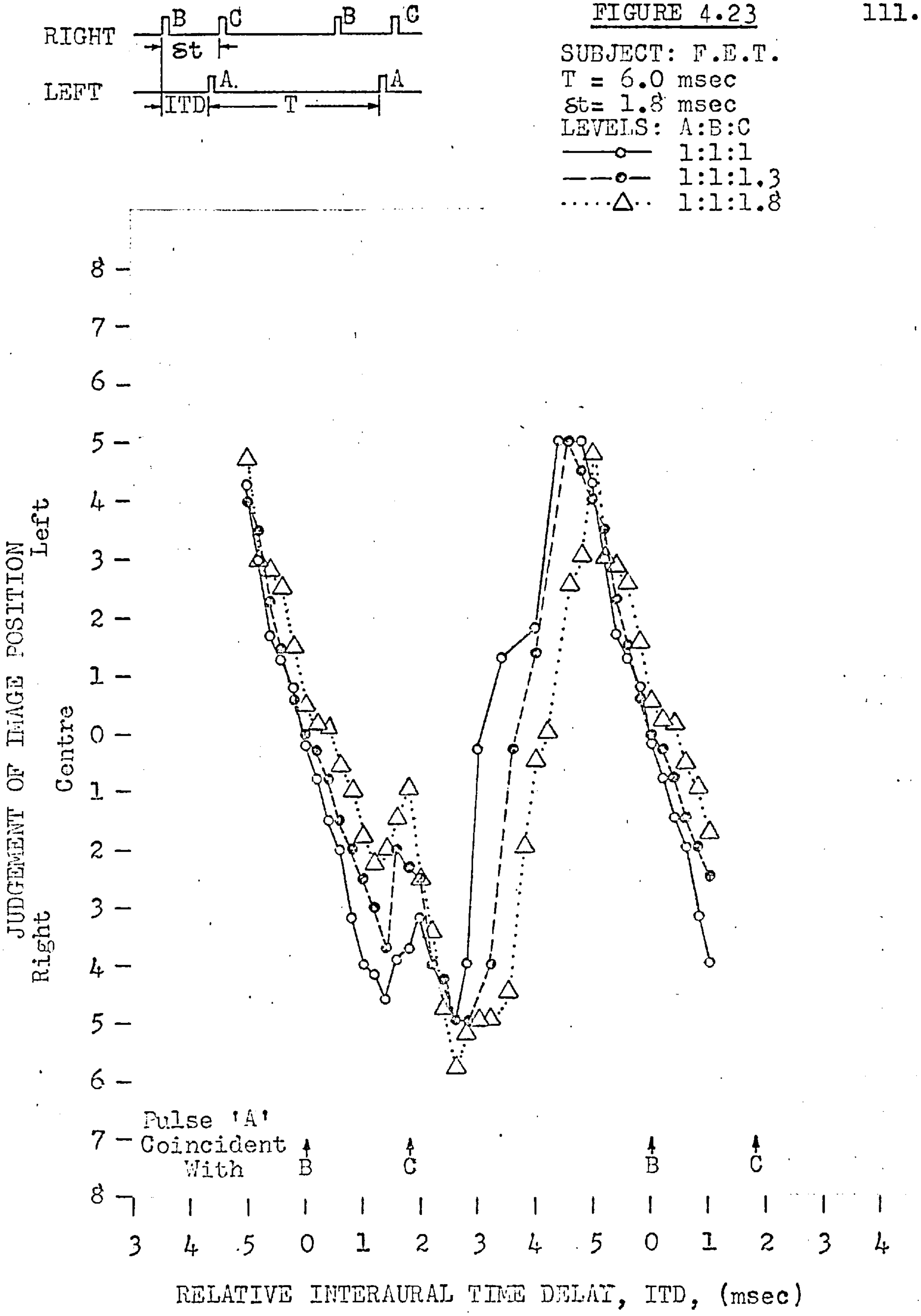
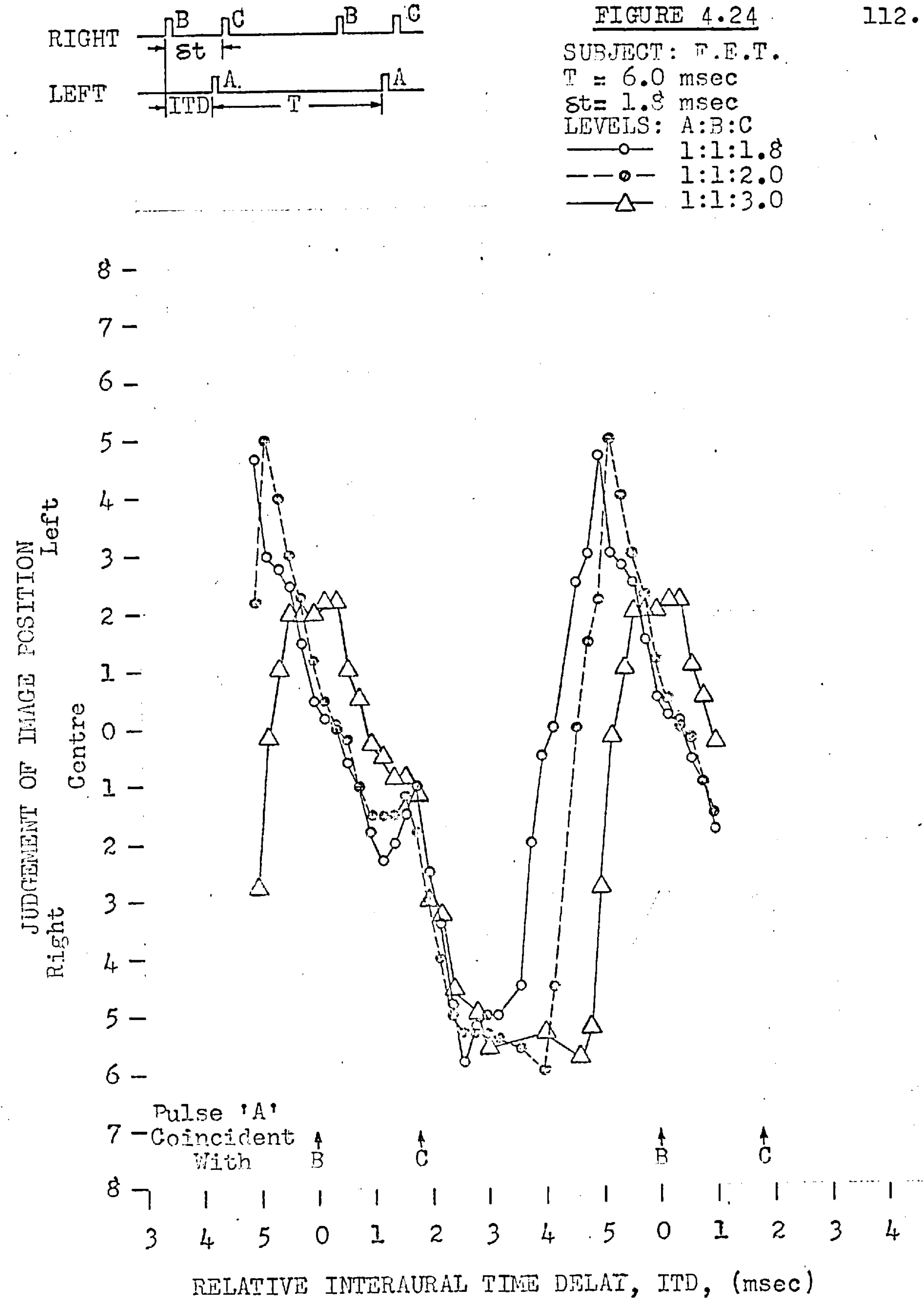

RIGHT
B
C
B
C
δt
LEFT
A.
A
ITD
T
SUBJECT: F.E.T.
T = 6.0 msec
δt= 1.8 msec
LEVELS: A:B:C
—○— 1:1:1
– –●– 1:1:1.3
····△·· 1:1:1.8
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
C
B
C
3
4
.5
0
1
2
3
4
.5
0
1
2
3
4
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 4.24



SUBJECT: F.E.T.
T = 6.0 msec
δt = 1.8 msec
LEVELS: A:B:C
—o— 1:1:1.8
- -⊘- - 1:1:2.0
—△— 1:1:3.0

FIGURE 4.25

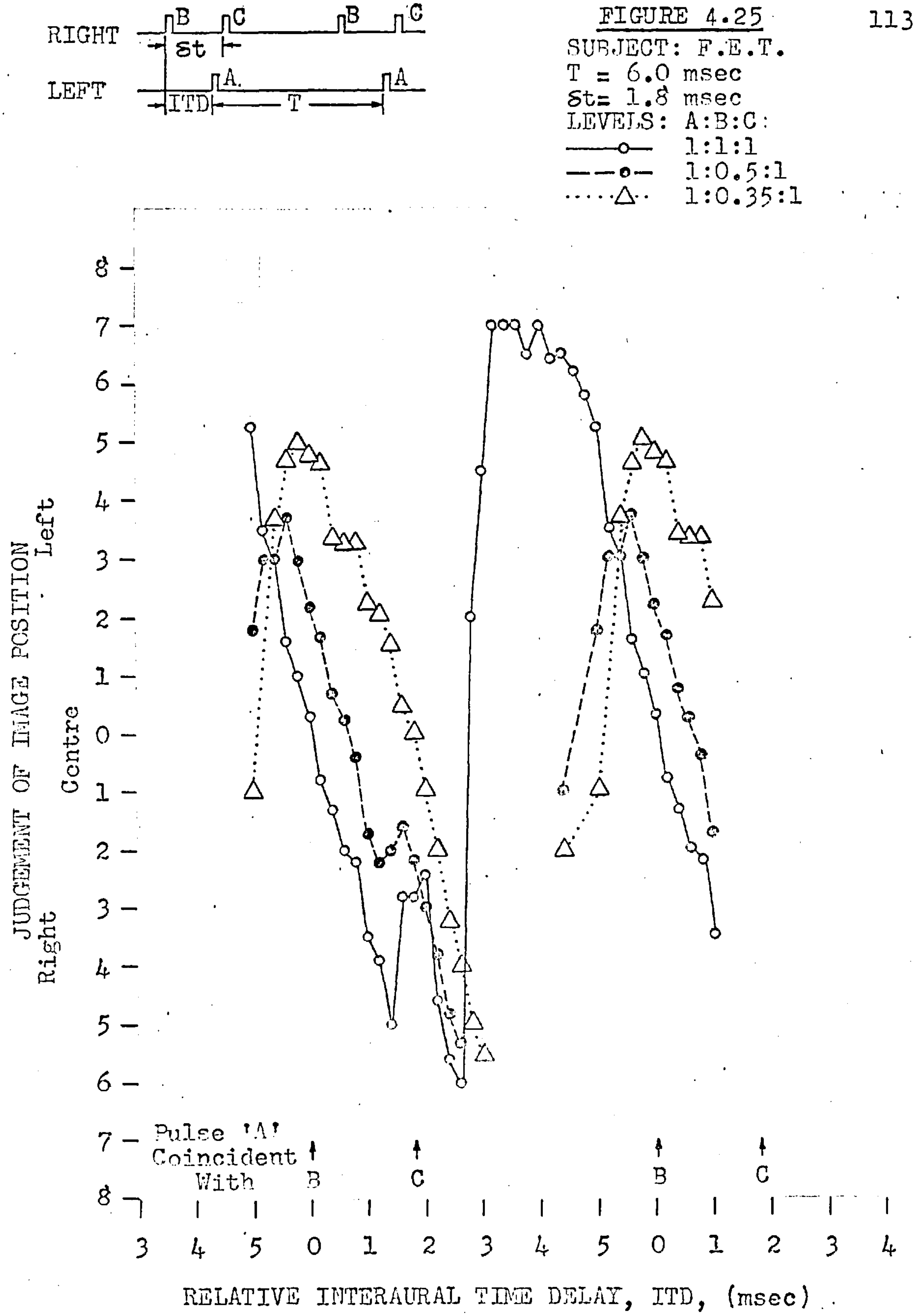
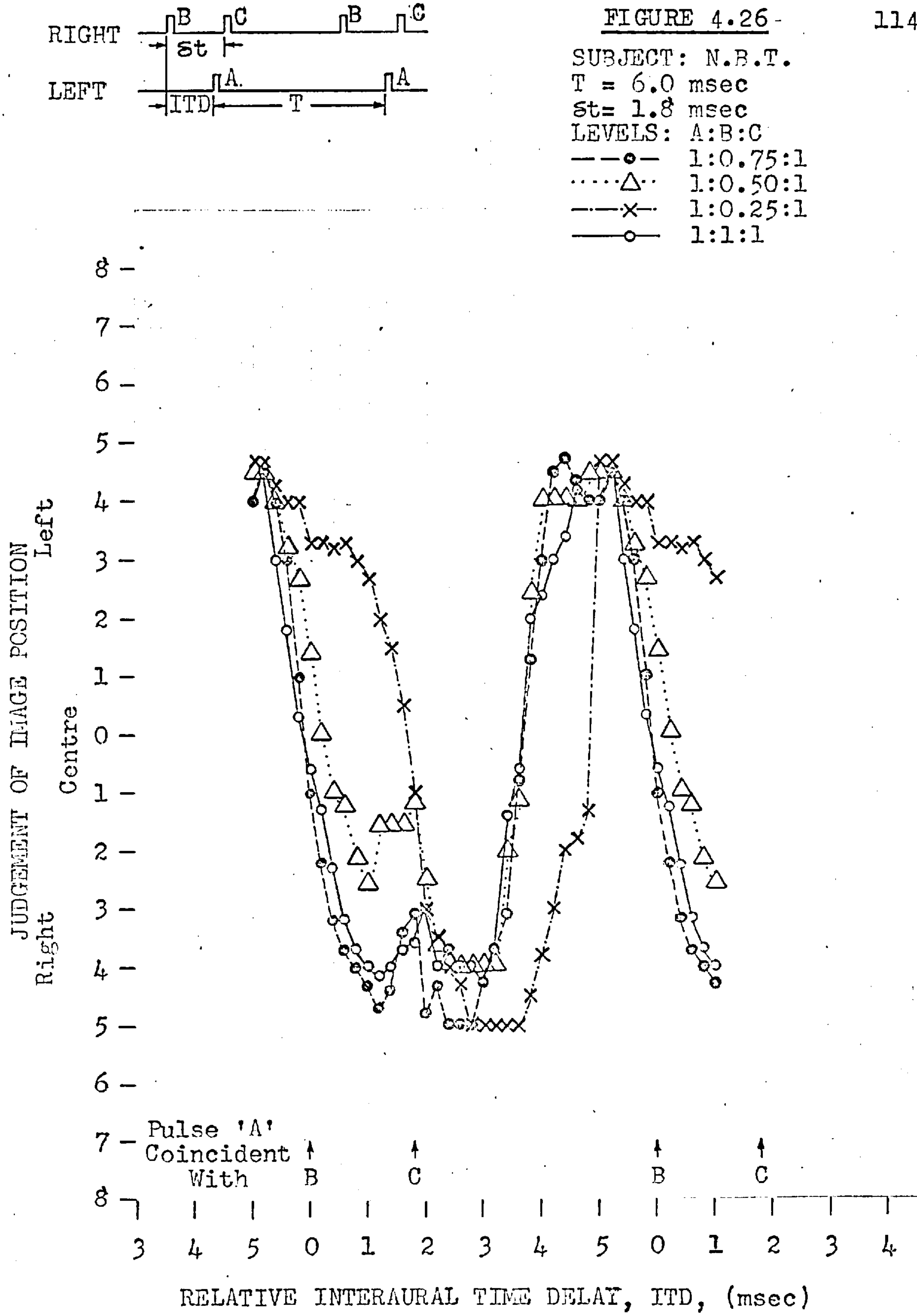

RIGHT
B
C
B
C
δt
LEFT
A.
A
ITD
T
SUBJECT: F.E.T.
T = 6.0 msec
δt= 1.8 msec
LEVELS: A:B:C:
1:1:1
1:0.5:1
1:0.35:1
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
C
B
C
3
4
5
0
1
2
3
4
5
0
1
2
3
4
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 4.26



SUBJECT: N.B.T.
T = 6.0 msec
δt= 1.8 msec
LEVELS: A:B:C

FIGURE 4.27

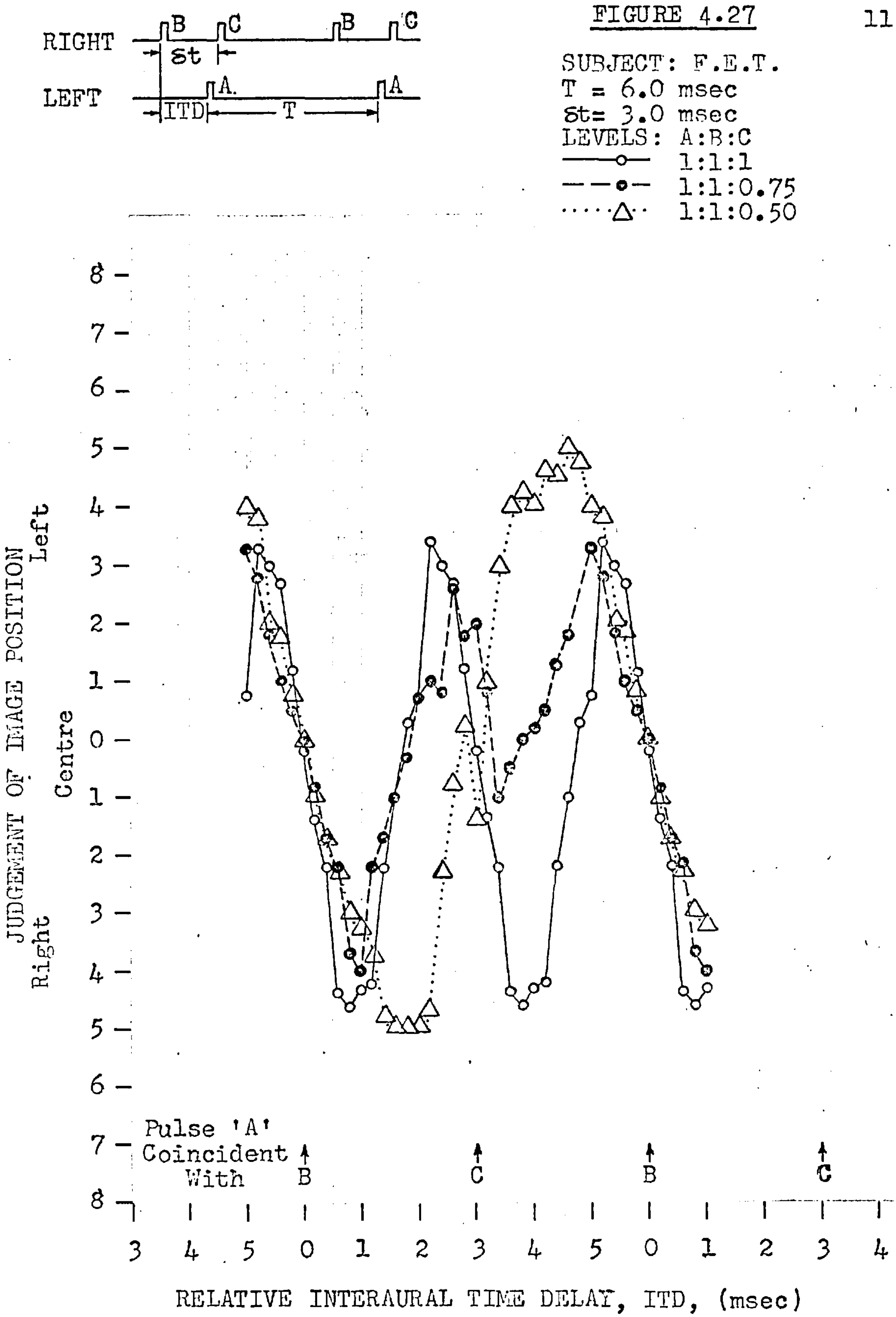

RIGHT
B
C
B
C
δt
LEFT
A.
A
ITD
T
SUBJECT: F.E.T.
T = 6.0 msec
δt= 3.0 msec
LEVELS: A:B:C
—○— 1:1:1
– –●– 1:1:0.75
····△·· 1:1:0.50
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
C
B
C
3
4
5
0
1
2
3
4
5
0
1
2
3
4
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 4.28

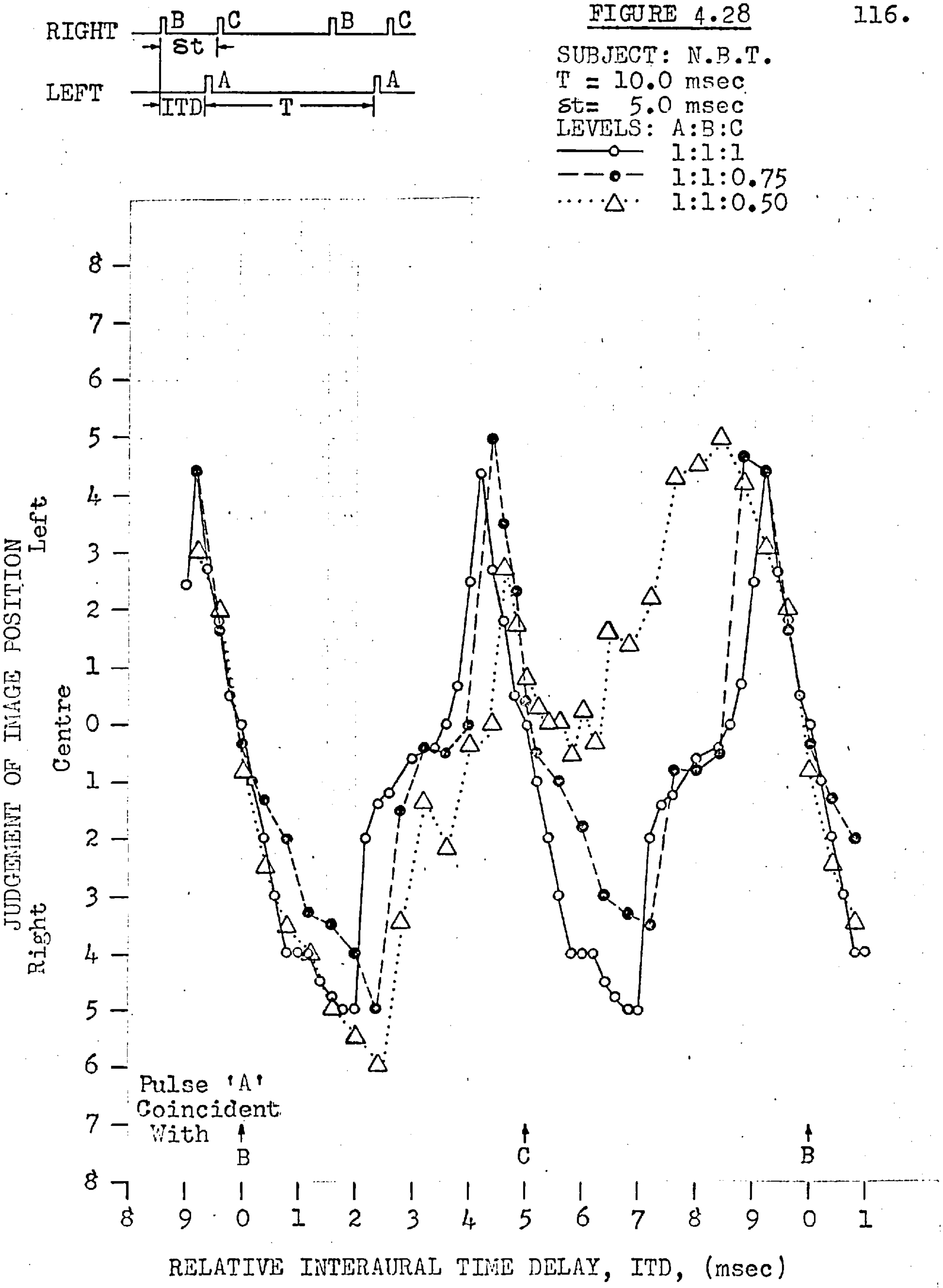

RIGHT
B
C
B
C
δt
LEFT
A
A
ITD
T
SUBJECT: N.B.T.
T = 10.0 msec
δt = 5.0 msec
LEVELS: A:B:C
1:1:1
1:1:0.75
1:1:0.50
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
C
B
8
9
0
1
2
3
4
5
6
7
8
9
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 4.29

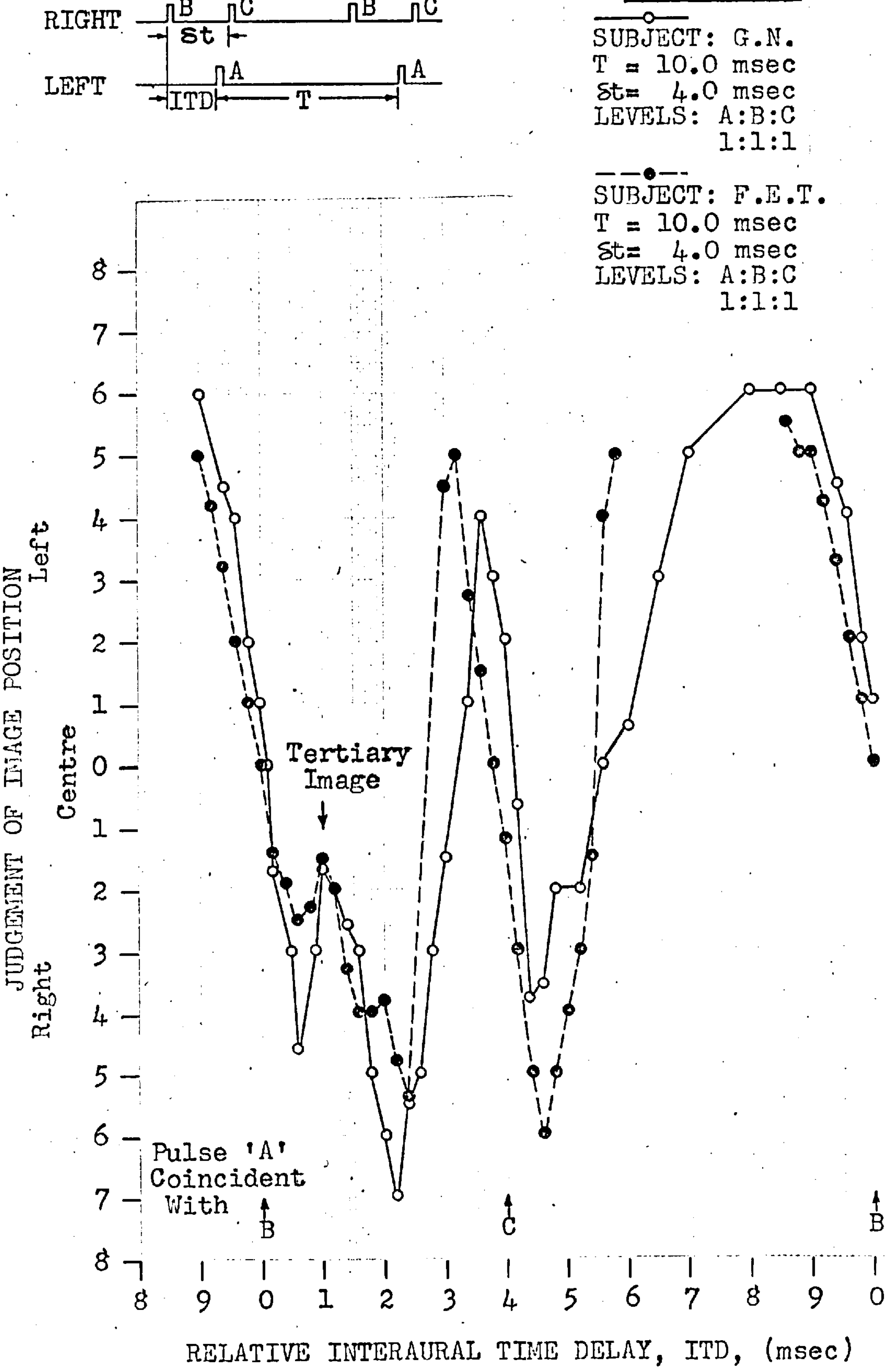

RIGHT
B
C
B
C
St
LEFT
A
A
ITD
T
SUBJECT: G.N.
T = 10.0 msec
St= 4.0 msec
LEVELS: A:B:C
1:1:1
SUBJECT: F.E.T.
T = 10.0 msec
St= 4.0 msec
LEVELS: A:B:C
1:1:1
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Tertiary Image
Pulse 'A' Coincident With
B
C
B
8
9
0
1
2
3
4
5
6
7
8
9
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 4.30

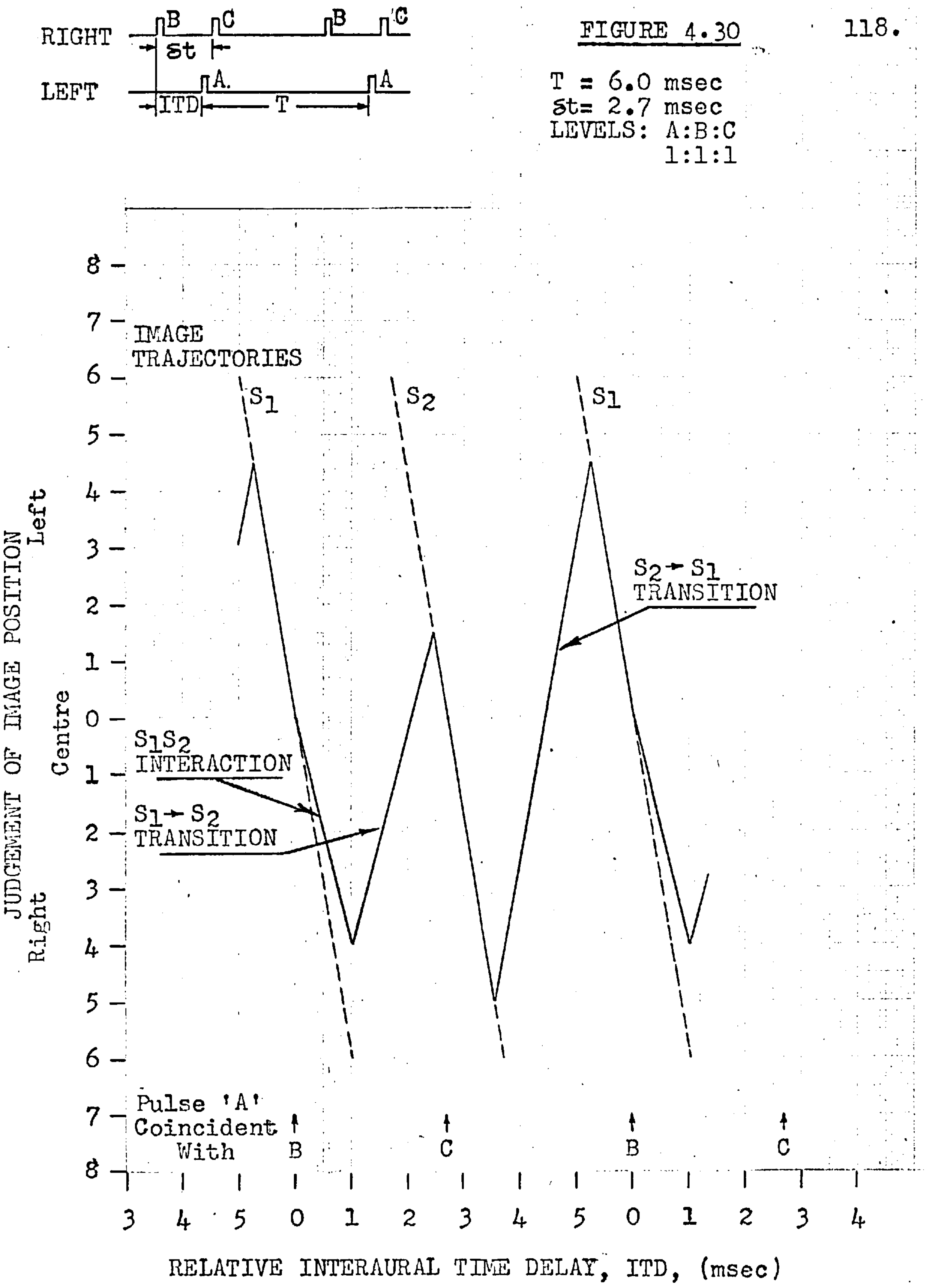
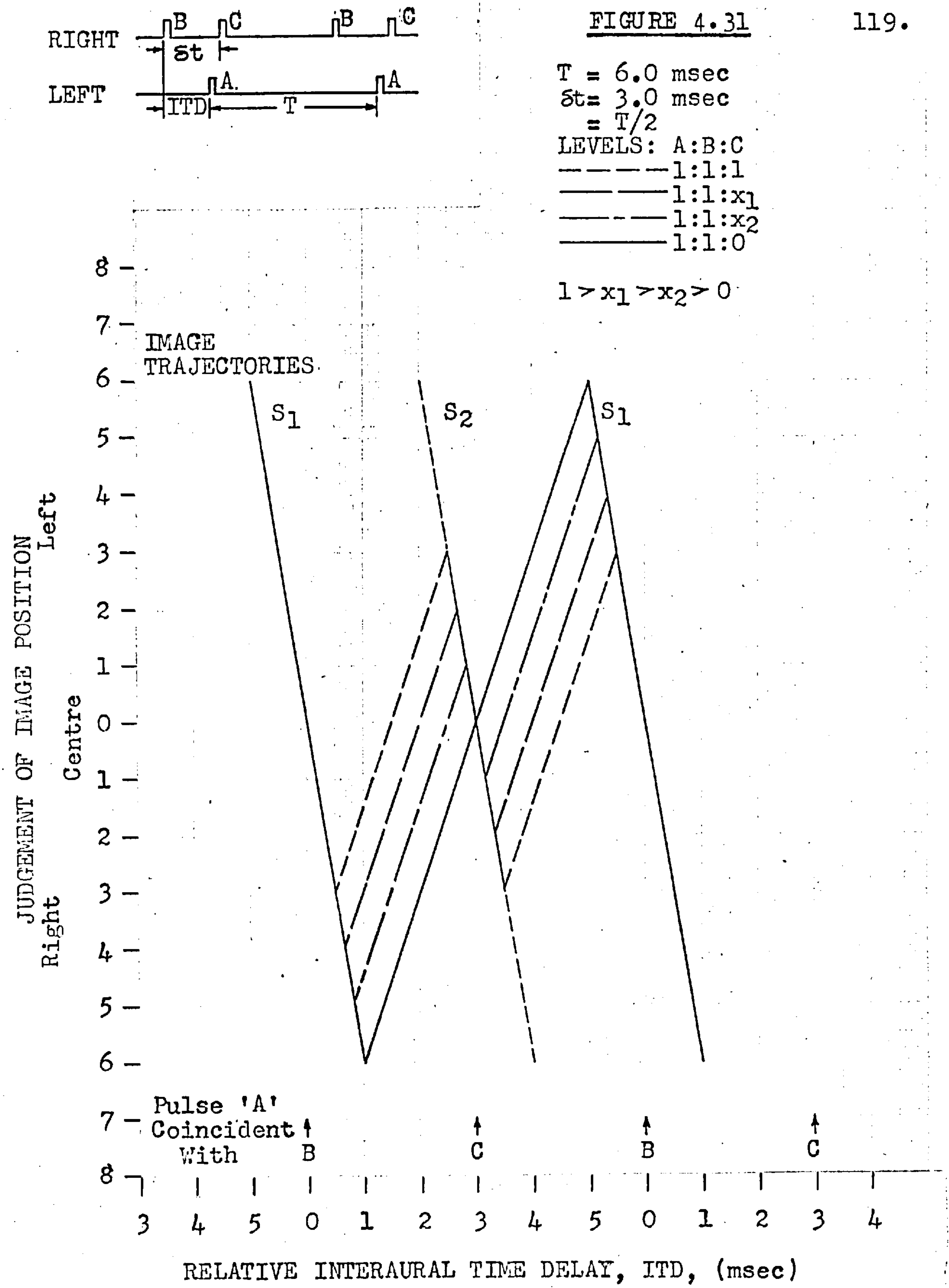
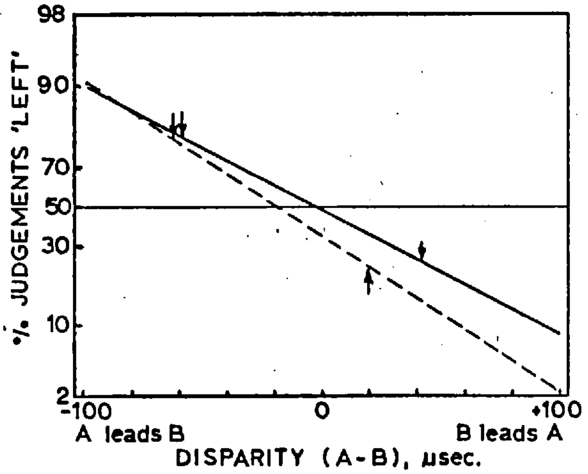
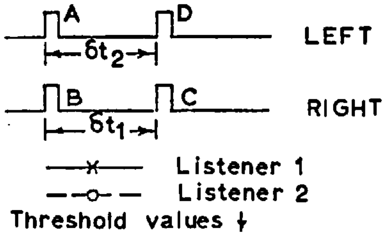

RIGHT
B
C
B
C
δt
LEFT
A.
A
ITD
T
T = 6.0 msec
δt= 2.7 msec
LEVELS: A:B:C
1:1:1
IMAGE TRAJECTORIES
S_1
S_2
S_1
$S_2 \rightarrow S_1$ TRANSITION
S_1S_2 INTERACTION
$S_1 \rightarrow S_2$ TRANSITION
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
C
B
C
3
4
5
0
1
2
3
4
5
0
1
2
3
4
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 4.31

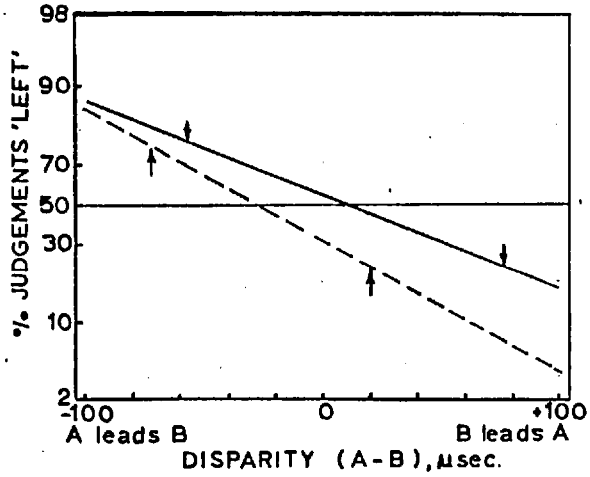

RIGHT
B
C
B
C
δt
LEFT
A.
A
ITD
T
T = 6.0 msec
δt= 3.0 msec
= T/2
LEVELS: A:B:C
1:1:1
1:1:x1
1:1:x2
1:1:0
1 > x1 > x2 > 0
IMAGE TRAJECTORIES
S1
S2
S1
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
Pulse 'A' Coincident With
B
C
B
C
3
4
5
0
1
2
3
4
5
0
1
2
3
4
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)





SIGNAL: A:B = 1:1

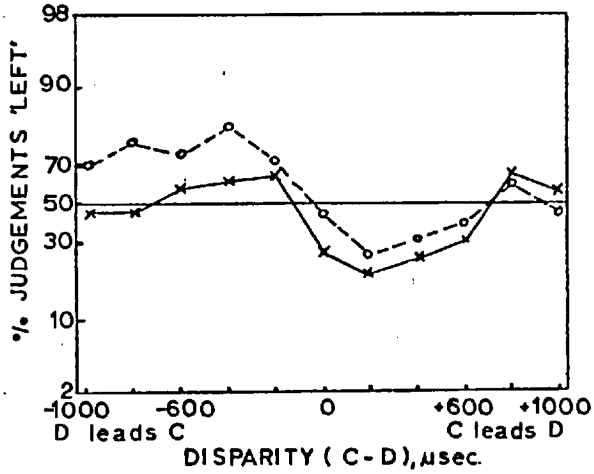
From Wallach, et al., 1949, Fig. 6.



(a)

DISPARITY C-D = 0
SIGNAL: A:B:C:D = 1:1:1:1
$\delta t_1 \simeq \delta t_2 \simeq 2.0$ msec

From Wallach, et al., 1949, Fig. 7.



(b)

SIGNAL: A:B:C:D = 1:1:1:1
DISPARITY A-B = 0
$\delta t_1 \simeq \delta t_2 \simeq 2.0$ msec

From Wallach, et al., 1949, Fig. 13

FIGURE 4.32

CHAPTER 5
MULTIPLE IMAGE LATERALIZATIONS

5.1.0 Introduction

The previous chapter has considered in some detail the lateralization characteristics of the dominant image perceived with binaurally presented acoustic transients. While, in general, the results were amenable to fairly direct interpretation, there were certain features of the experiments which warranted further elucidation, viz., the occasion of multiple image reportings by certain listeners and the observation by some of the more experienced listeners that there was a change in the calibre, character or number of images perceived as ITD was altered. The purpose of this series of experiments was to seek in an objective fashion the essence of these changes.

Images occurring simultaneously may have broadly similar characteristics (as, for example, in the transitional region between S_1 and S_2 or S_2 and S_1 image trajectories) or may have specific individual characteristics (as, for example, in the images of a binaural two-component tone). The former multiplicity of images will be referred to as "split images", and the latter as "multiple images." The emphasis in these experiments was on the latter. Multiple images were lateralized in a variety of situations, from the simple case of binaural two-component tones where the perceived images are of tonal character having, perhaps, quite different pitch characteristics, to the much more complex case of binaural repetitive transients where an array of images may be identified and independently lateralized. In the latter situation the crucial question is whether any or all of the multiple images arise, perhaps solely, as images of the separate harmonic components of the repetitive transient signals.

5.2.0 Experimental Technique

The technique applied in these experiments is essentially that described in section 4.3.0. The important innovation of the present experiments is that listeners were required to identify and independently lateralize a number of images at each setting of the ITD parameter.

The first set of experiments employed, as a signal, binaural, equal level, two-component tones. Here, the ITD between each of the component tones was adjusted in equal increments, and positional judgements pertaining to the two tonal images were recorded at each randomized setting. ITD values ranging over a number of periods of the sinusoids were employed. Listeners had little difficulty in reliably identifying the two images arising from binaural two-component tones in the frequency range 200 cps to 1500 cps; only when the frequency separation between the components was less than about 50-100 cps did the task become more difficult, and for some less experienced listeners, impossible.

The remainder of the experiments employed binaural single- and double-pulse trains as were previously used. A few listeners in the previous experiments spontaneously reported hearing other images along with the normally dominant impulsive image; indeed, the impulsive image itself was perceived by some to be comprised of at least two components having different pitch-like properties. The listener's task in these experiments was, therefore, to identify and report positions of these various images.

It will be seen in the forthcoming results that judgement transitions in these curves are generally much sharper than many of those in previous measurements. This has resulted from the innovation of requiring the listener to report split images in their apparent order of importance. These ranked judgements were then plotted separately. Most of the curves show, for simplicity, only the dominant elements of the split-images where these arise; others, where clarity of the presentation is assisted, show both primary and secondary elements of the reported split images. Listeners occasionally reported images as being somewhat vague or dispersed; this information was recorded, and is indicated on the graphs by symbols with superimposed broad vertical lines.

The experimental results which led to these modifications of technique were the consequence of critical listening on the part of one subject, whose listening experience was, by this stage, considerable. This listener quite spontaneously recognized that there were present, under some conditions, images of apparently different importance and character. In particular, where split-images were identified it seemed that in most instances one of the elements was in some sense more important than the other. Also, on certain occasions there seemed to be a third image present, having a less clearly impulsive, almost tonal, character.

On this evidence a few pilot experiments were arranged in which this listener reported, in order, apparent positions of the dominant impulsive image, the secondary impulsive image, and the tertiary image. Results of the first experiment of this type are illustrated in Figure 5.1, where the results are shown plotted in the style of the earlier experiments, i.e., averaging all judgements at each ITD value, and also when judgements are plotted separately according to the individual categories in which they were reported. It was apparent from the judgement groupings (supported by the listener's introspective comments) that the first and second judgements, where both were reported, pertained to images which were positioned at the lateral extremes. These transitions were, therefore, apparently indicative of a shift of attention from one image to the other, and not of two images tracing out opposing trajectories as ITD was increased. It is relevant in this respect to refer to the results of Sayers (1964) and those of the previous chapter where listeners showed what could also be described as a transference of attention between the elements of a split-image. The judged trajectory of the tertiary image was seen to be positioned midway between S_1 image trajectories, and appeared to have a similar slope.

Upon further critical listening in casual experiments this subject reported that the tertiary image could be identified at all values of ITD, and that it appeared to follow the course of the main impulsive image during its centre crossing near ITD = 0 thus crossing the centre axis twice per pulse repetition period. Moreoever, there appeared to be other images which could be repeatedly identified. In the results of a formal experiment where these new images were independently identified and lateralized it was seen that certain of the images crossed the centre axis more than twice within a span of ITD equal to the repetition period. It was determined that these new images had pitches which could be matched to tonal signals at frequencies which were harmonics of the pulse repetition frequency. To assist in the identification of these tonal images in subsequent experiments, reference tones at these frequencies were provided which, at the listener's requirement, would replace the experimental signal for a brief interval. With this means of directing the listener's attention to the various harmonic frequencies, other listeners not able to identify tonal images in earlier experiments, could now easily and reproducibly lateralize these new images. With a small amount of practice, only occasional reference to the tonal signals was required. Unfortunately, no such simple artifice was available in the case of the impulsive images.

Most listeners, in the experiments reported here, were able to lateralize tonal images associated with frequencies up to the fifth or sixth harmonic of the pulse repetition frequency. When, however, the repetition period exceeded about 12 msec, or the harmonic frequency exceeded about 1500 cps, tonal images were exceedingly difficult to identify. For the acoustic transient signals and the repetition periods employed in these experiments (0.1 msec pulses applied to Sharpe HA-10 earphones; T = 6-12 msec; signal level = 35-40 dB SL) it was very difficult to identify the tonal fundamental. This situation could be alleviated either by increasing the pulse width or by substantially increasing the signal level. The former was undesirable for reasons of continuity with previous experiments, and the latter was abandoned because of the intolerable signal level required. Therefore, the tonal fundamental image was not lateralized in the reported experiments, although in the appropriate circumstances it was easily identified and predictably lateralized.

These experiments often required that listeners should identify and lateralize as many as six coexisting images. Needless to say, such experiments were very fatiguing as the number of such image identifications and positional judgements was, of necessity, extremely large (up to 1200 in one experiment). In many cases, therefore, the results reflect only two or three judgements for each experimental condition.

5.3.0 Apparatus

In the experiments with two-component tones each of the component tones was derived from an oscillator (Dawe, No. 445A) with associated variable-phase adapter (Dawe, No. 446A); the fixed-phase outputs were applied to the inputs of the summing amplifier for one channel and the variable-phase outputs were applied to the summing amplifier for the opposite channel. Suitable adjustment of the phase shifts of the component tones in one channel introduced an effective ITD between the left - and right-channel signals. The remainder of the apparatus was as illustrated in Figure 4.1.

Experiments using binaural transients utilized an arrangement of equipment as is illustrated in Fig. 4.1. Slight modification of this scheme was required for experiments in which the IAD of one harmonic component was altered. This was accomplished by the addition, to one channel, of a suitably phased and phase-locked sinusoid at the harmonic frequency. The tonal signal was derived from an oscillator and associated variable-phase adapter. The oscillator was set to the appropriate harmonic frequency and the fixed-phase output used as a synchronizing signal for the Clock and Delay 1 generator; the variable phase output was passed through an attenuator and then applied to the spare input of the summing amplifier of the right-hand channel. Accurate phasing of the added harmonic tone was accomplished by adjustment of the variable-phase adapter.

Spectral amplitude measurements, where stated, were obtained by analysis of the acoustic signal obtained with the artificial ear (Bruel & Kjaer type 4109, with flat-plate coupler) and a wave analyser type FRA2b by Radiometer.

5.4.0 Experiments

5.4.1 Multiple Images with Binaural Two-Component Tones

In these experiments listeners were required to identify and report positions of the separate images associated with each of the two binaural tones. All listeners had little difficulty with this task when the frequency separation of the tones were substantial; when, however, the separation was reduced to less than about 100 cps, there was a considerable beating effect and many listeners had difficulty in identifying the individual components. In all cases where the image identification was effected, the results were well described for each component by curves of positional judgements versus ITD for binaural single-component tones at the same frequency.

Fig. 5.2 illustrates results of an experiment in which the frequency separation was only 20 cps ($f_1 =$ 600 cps, $f_2 = 620$ cps). This is the least separation at which listeners in this study were able to independently lateralize the two tonal images. It is interesting to note that the width of the critical bands at these frequencies is about 100 cps (Zwicker, et al., 1957).

5.4.2 Multiple Images with Binaural Transients: Tonal Images

In these experiments using binaural trains of single or double acoustic transients listeners were required to identify some or all of the readily perceptible tonal harmonic images as well as the normally dominant image of impulsive character. Where the listener felt he could reliably identify more than one image of impulsive character he was encouraged to do so. Special consideration will be given to the impulsive images in the following section, and results pertaining to these will be illustrated here along with tonal image results where both were acquired in the same experiment.

The curves of Figs. 5.3, 5.4 and 5.5 illustrate for binaural equal-level pulse trains, positional judgements with ITD for two impulsive images, one reported as being of low-pitched character, another reported as being of high-pitched character, and for the harmonic tonal images from the second to the fifth. It is clear from the periodicities in ITD, of the judgement trajectories that the listener was indeed attending to, and lateralizing images arising from binaural interaction of the individual harmonics of the periodic pulse signal.

It was immediately interesting to investigate whether the harmonic images could be manipulated by various methods. The simplest of these manipulations was the introduction of an overall IAD by an increase of the relative pulse amplitude on one side. Fig. 5.6 illustrates, for the representative second and fifth harmonics the effects of such an IAD introduced to favour first one side and then the other; for comparison the IAD = 0 situation is included. The results shown here were obtained in an interlaced fashion. An overall signal IAD would be expected to increase each of the harmonics on one side by a similar amount. The vertical shifts of the judgement curves, which may therefore be anticipated, may be seen for each of the harmonic image trajectories shown.

The second method of manipulation employed a change in the IAD of one harmonic component only. This was accomplished by the addition of a suitably-phased, frequency locked, sinusoid at the same frequency, to the repetitive transient signal at one ear. The phase required for the added sinusoid so as to sum exactly in phase with the component already present was deduced initially by experimental observation. Replacement of the transient signal in one channel by a single phase-locked sinusoid at the appropriate harmonic frequency gave rise to a single tonal image. Lateralization judgements of this image indicated centre crossings which would be characteristic of binaural interaction of the experimental tone with a contralateral tone co-sinusoidally related to the periodic pulse train, i.e., as expected the harmonic is in a cosinusoidal relation to the brief pulses. Subsequent application of this conclusion to results of experiments using binaural pulse trains indicated satisfactory agreement.

Fig. 5.7 illustrates results of an experiment of this type; the second and fifth harmonics have again been chosen as representative. Here, the second harmonic component in the right-ear signal was enhanced; a tone at the second harmonic frequency was added at an arbitrary reference level, and then at 10 dB above this level. Again, the IAD = 0 results are included for comparison. Positional judgements for the second harmonic image were, as expected, progressively displaced to the right as IAD was increased; the fifth and other harmonic trajectories showed only small, apparently normal fluctuations with no significant lateral displacement.

Another method utilizes a change in the temporal pattern of pulses in one signal; it is possible, for example, to interpolate an extra pulse within each period of the transient signal. The spectra of the resulting periodic train of double-pulses are manipulable by adjusting the spacing (δt) between the members of the pulse pair, or by altering the relative amplitudes and polarities of the B and C pulses.

Figs. 5.8, 5.9 and 5.10 illustrate judgement trajectories of the low and high pitched impulsive images and of the second to the fifth harmonic images for the situation of periodic single pulses in the left earphone and periodic double pulses ($\delta t = 1.8$ msec) in the right earphone. For comparison, curves for an unmodified right-ear signal are included in the diagrams; these results were taken in the same experiment with the same listener.

It is clear that the interpolation of an extra pulse per period to the right-hand signal will alter the amplitudes and phases of some or all of the harmonic components in that signal. The amplitude spectra of the left-and right-ear signals were measured and the results are stated on the diagrams as effective IAD's for each harmonic. The phase spectra are less easily obtained. A first approximation, however, may be had by assuming that the phase shifts of the lower harmonics due to the audio amplifiers and earphones are so small as to be neglected; the favourable comparison of the measured acoustic transient and the electrical signal pulse supports such a view. Assuming also that the pulse width is small compared to the harmonic periods, the single-pulse trains may then be represented for the low frequencies in question by an harmonically related series of equal-level cosinusoids. The double-pulse train, considered as two single-pulse trains separated by a time delay (δt), then has a phase spectrum calculable as the sum of two equal-level cosine spectra separated in time by δt.

The interaural phase of each binaural harmonic component determines the ITD at which the judgement trajectories would be expected to cross centre. For interest, the points of expected centre crossing (ITD= 0 condition) of the harmonic image judgement trajectories are shown in the figures. These predictive points do not allow for the effects of IAD on the positional judgements; in most instances such effects would be expected to be small. It will be seen that agreement with the results is generally quite good.

The substantial phase shift of the second harmonic component of the right-ear signal is well demonstrated by a shift of the judgement curve along the ITD axis in Fig. 5.9. The judgement curve of the fifth harmonic image (Fig. 5.10) is seen to be displaced well to the left; such a bias would be anticipated on account of the large IAD of that spectral component. It was difficult, because of the ambient acoustical and electrical noise, to make measurements of harmonic amplitudes at low levels; however, in this case the amplitude of the fifth harmonic was certainly reduced by more than 30 dB by the introduction of the second pulse per period.

Inversion of either of the B or C pulses produces another form of spectral modification. Figures 5.11, 5.12 and 5.13 illustrate superimposed judgement curves for two experiments conducted in interlaced fashion, in which the C pulses were inverted; the A pulses in one case were of normal (condensation) polarity and in the other were inverted. Judgement trajectories for two impulsive images and for the second to the fifth harmonic images are shown. Again, the centre-crossings of the image trajectories were closely predicted by the simple method applied in the previous experiments.

The last manipulation of this type to consider is a change of the relative amplitudes of the B and C pulses. The curves of Fig. 5.14 illustrate judgement trajectories for the second and third harmonic images where the level of the C pulses was increased. The effect is again as would be expected from consideration of the signal spectra.

5.4.3 Multiple Images with Binaural Transients: Impulsive Images

In each of the experiments reported in the previous chapter listeners were required to report the position of the dominant image. This image was generally reported to be of impulsive character. Some listeners spontaneously communicated their impressions of two simultaneously perceived impulsive images: the normally-dominant one of low-pitched character and one of high-pitched character.

From results of certain experiments where listeners gave lateralization judgements for these impulsive images it was clear that the judged image trajectories have important differences. Fig. 5.8, for example, illustrates judgement trajectories for the case of repetitive single versus double transients. It is evident that the high-pitched impulsive image follows almost entirely the expected S_1 image trajectory with little apparent influence by the S_2 image; however, the judged trajectory of the low-pitched impulsive image, the normally dominant image, shows the characteristic effects of the S_2 image (see also Fig. 4.7).

Comparison of these curves with those obtained using binaural repetitive single transients (illustrated on the same diagram) reveals a few interesting differences.

Consider first the low-pitched impulsive image trajectories. The introduction of the second pulse (C) per period has had the effect of displacing the major judgement transition from midway between S_1 centerings to about midway between S_2 and S_1 centerings. This is essentially what would be predicted from the simple argument wherein the S_1 and S_2 images are hypothesized to coexist and judgements transition from one to the other according, inter alia, to their relative importances and spatial positions. In the case of the high-pitched impulsive image, however, the presence of the S_2 image was not apparent; in fact, the major judgement transition occurs in the region where the S_2 image trajectory is hypothesized to exist and here this transition would be least expected. Indeed, the observed trend is suggestive of an IAD favouring the left side, and an S_2 image which is either entirely absent or greatly diminished in importance.

The curves of Fig. 5.11 further illustrate this phenomenon and demonstrate similar features.

5.5.0 Comparative Trajectories of Tonal and Impulsive Images

Repetitive binaural transients may lead to the simultaneous perception of two types of images; the normally dominant impulsive images and various harmonically related tonal images. It is interesting to enquire if the commonly dominant low-pitched impulsive image arises as a synthesis in any sense of the independent harmonic images.

A definitive experiment in this context would use binaural transients from which impulsive and tonal images arise and in which numbers of the tonal images do not move concurrently across the auditory space as ITD is altered.

Experiments using repetitive single versus double transients are of this type. It has been demonstrated that the addition of an extra pulse per period to the pulse train driving one earphone modified the interaural amplitude and phase of the individual harmonic components; where this was so, the resulting trajectories of the harmonic images were seen to have been displaced along the ITD axis. Different pulse-pair spacings (δt) would give rise to different harmonic phase shifts, and therefore to different centre-position crossing values of ITD for the various harmonic image trajectories. Figs. 4.4 to 4.17 and Fig. 5.8 illustrate for a variety of pulse-pair spacings and repetition periods that the trajectory of the main impulsive image remains undisplaced-crossing centre in each case at closely ITD = 0. These observations are confirmed in the results illustrated in Fig. 5.11 where the pulse polarities were manipulated. The comparative curves shown are enlightening in a slightly different respect, for between the two curves illustrated, the interaural phases of the individual components have been, in each case, altered by precisely π.

As a further verification of this latter point, the simple cases of cophasic and antiphasic binaural single repetitive transients were studied. Fig. 5.15 shows judgement trajectories for the low - and high-pitched impulsive images for these two situations with the same listener. These results were obtained in the same experiment. Impulsive image trajectories exhibited no significant displacement in respect of centre-crossing values of ITD, while the harmonic image trajectories (not shown) were, as in Figs. 5.9 and 5.10, each displaced by precisely half the harmonic period along the

ITD axis.

It would therefore seem that the impulsive images do not arise as a simple synthesis of independent harmonic images.

5.6.0 Discussion

In the case of binaural two-component tones of low frequencies, listeners were able to identify and independently lateralize images arising from binaural interaction of each component, even when the frequency difference was as low as 20 cps at which separation the beating effect was considerable. That this was possible was interesting in the light of the published magnitudes of critical bands at the frequencies used - approximately 100 cps at frequencies around 600 cps (Zwicker, et al., 1957). It may be relevant, however, that recent studies have demonstrated that the separability of monaurally presented multiple low-frequency tones is not only a function of frequency separation, but of the number of components present. Plomp (1964) found that the frequency differences between just-distinguishable partials of a complex (more than two components) tone closely correspond to the measured critical bands; however, two-component tones in the frequency range below about 1000 cps appeared to be more easily separated, showing a required minimum frequency difference of about 50 cps at 600 cps, and the still lower value of about 20 cps at 200 cps.

Binaural presentation of repetitive transients represents a situation where complex arrays of acoustic images may be simultaneously perceived. Listeners in these experiments have been able to independently track the coexisting harmonic and impulsive images with impressively reproducible results. Tonal harmonic image trajectories were demonstrated to be independently manipulable according to the relative interaural amplitudes and phases of the individual binaural harmonic components of the repetitive transients, and, indeed, were closely predicted on the basis of a simplified harmonic representation of the pulse trains.

Results of a number of experiments suggested that the impulsive images were not due to a synthesis, in any simple sense, of the low-frequency harmonic images. It seems that the impulsive images are apparently representative of some predominant waveform features of the acoustical signal and are possibly, therefore, established by a process different from that whereby the associated harmonic tonal images arise.

It is to be noted that the technique of independently tracking the various coexisting images and recognition of the importance of apparent transitions of attention between images according, inter alia, to their spatial disposition, are key factors in elucidating this involved situation. The simple procedure of requiring a listener to judge an unspecified or undifferentiated image in such experiments may yield results which are misleading; judgements may, at different values of ITD, relate directly to images of different character, or may reflect an apparent balance of attention between images of potentially different lateralization characteristics. This is an extension of the conclusions of the previous chapter, particularly those illustrated by the discussion of results by Wallach, et al., (1949).

While listeners generally had little difficulty in independently lateralizing the various harmonic images, not all listeners discerned the low-and high-pitched impulsive images spontaneously reported by some.

The dominant image was reported by all listeners to be of impulsive character. Where two impulsive images were identified the dominant one was commonly reported to be the image having the apparent low-pitched attribute. Support is given to this subjective impression by a comparison of low-pitched impulsive image trajectories with trajectories in comparable experiments of the single dominant image lateralized in experiments of the previous chapter (for example, Figs. 5.8 and 4.7).

The judged trajectories of the low-pitched impulsive image show, as did those for the dominant image in earlier experiments, effects apparently due to the presence of both S_1 and S_2 images. Judged trajectories of the high-pitched impulsive image, while bearing a certain resemblance to those for the low-pitched impulsive image, exhibited some rather different characteristics which are not readily explicable in the same terms. Clearly there was a case for further exploration of the impulsive images, with particular emphasis on the origins of these phenomena.

5.7.0 Conclusion

Results of a number of experiments have been described and discussed, where listeners have demonstrated the ability to identify and independently lateralize a number of coexistent images of tonal and impulsive character. In the case of binaural two-component tones, images were identified which related to each of the binaural components and these were lateralized as if each occurred alone. In the case of binaural repetitive transients, listeners were able to lateralize images of tonal character as well as the normally dominant impulsive image. Such images were well described as arising from binaural interaction ~~of~~

of the products of a spectral analysis of the repetitive transient signal in each ear. The impulsive images were held to arise from a process different from that whereby the tonal images arise, and appeared to be related to some important features of the acoustic waveform. There are, however, some crucial questions yet to be answered in respect of the impulsive images. It is the aim of the next chapter to deal with studies of these in some detail.

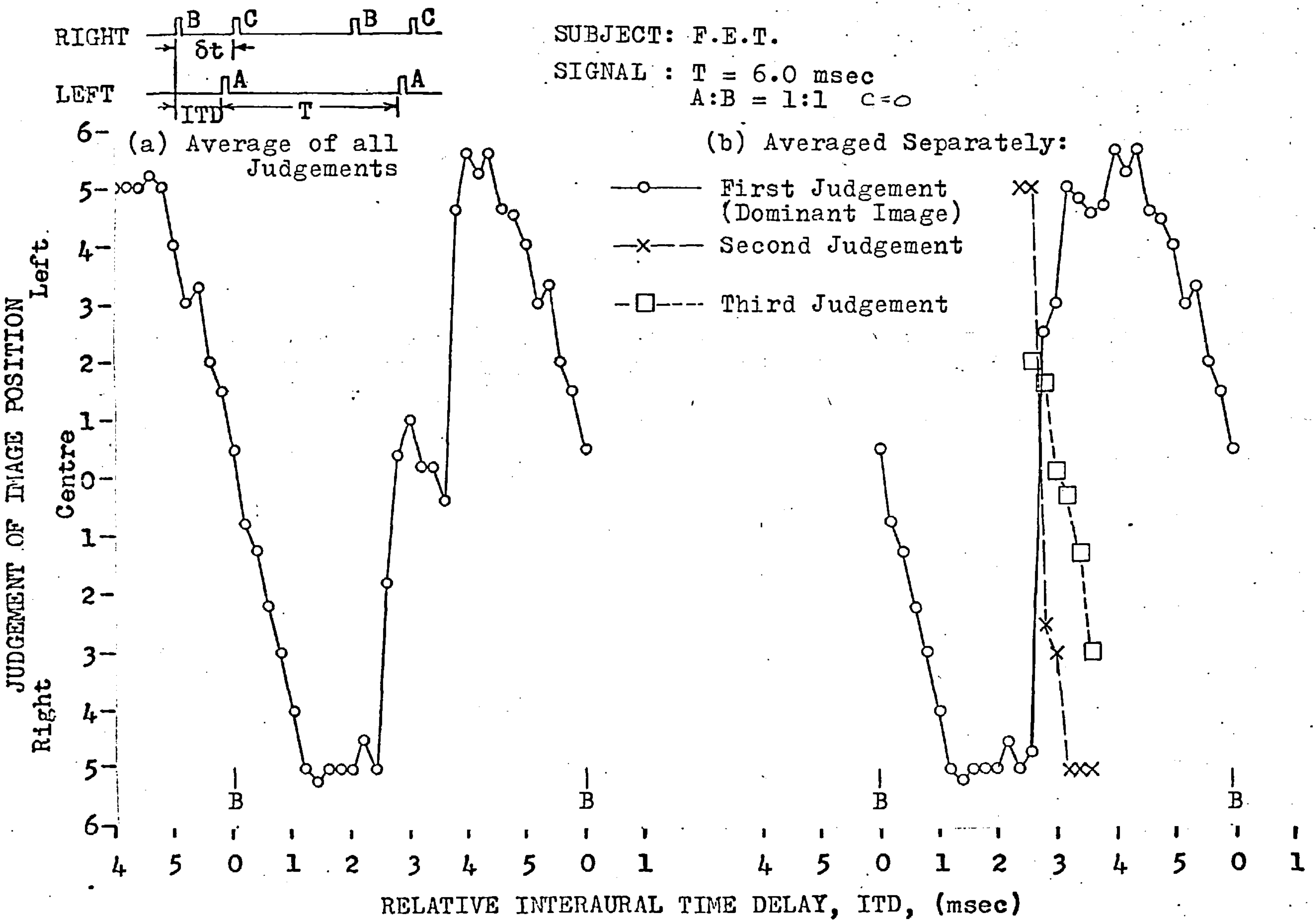

RIGHT
B
C
B
C
δt
LEFT
A
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec
A:B = 1:1 C=0
(a) Average of all Judgements
(b) Averaged Separately:
First Judgement (Dominant Image)
Second Judgement
Third Judgement
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
B
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.1

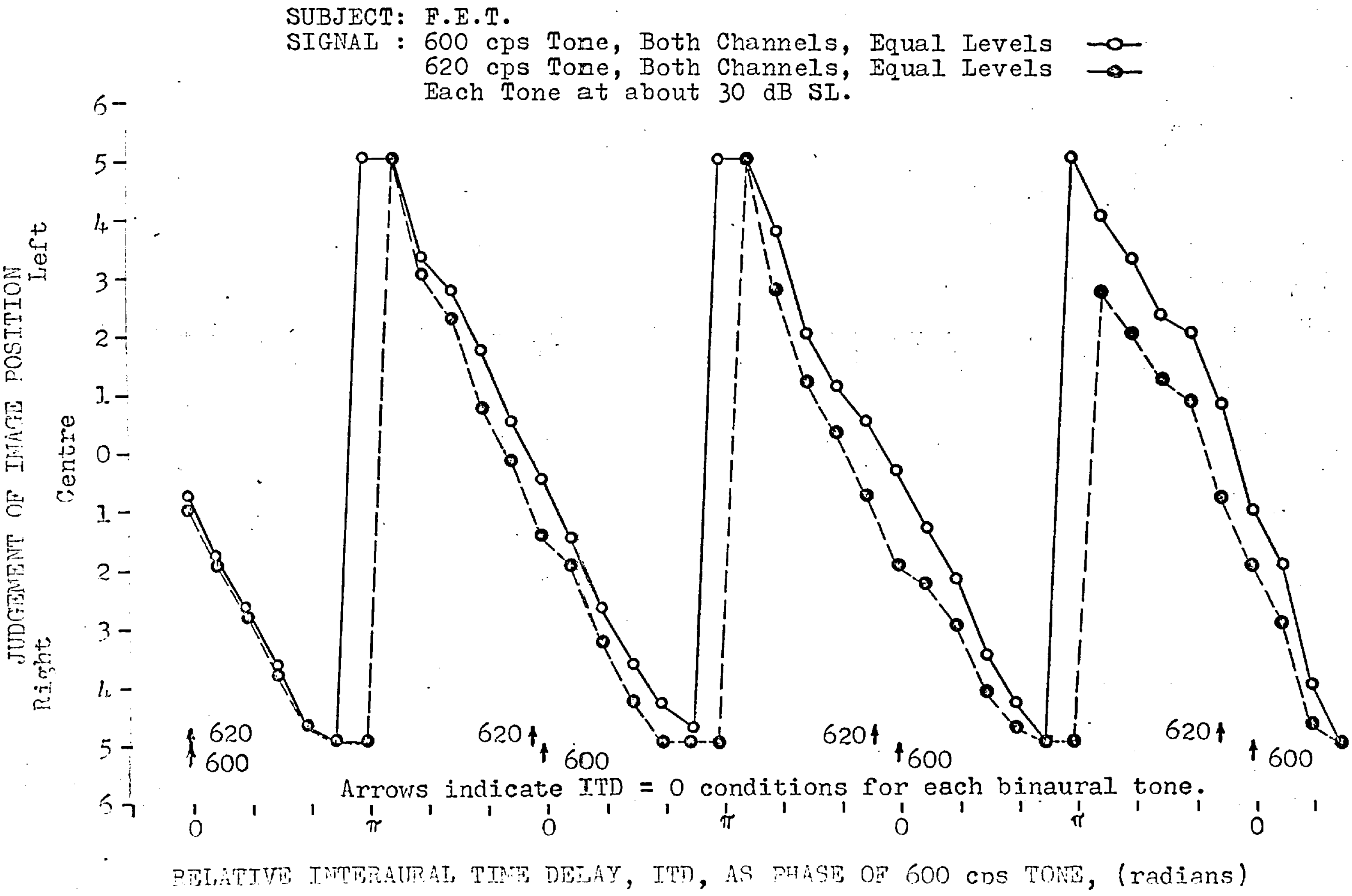

SUBJECT: F.E.T.
SIGNAL : 600 cps Tone, Both Channels, Equal Levels
620 cps Tone, Both Channels, Equal Levels
Each Tone at about 30 dB SL.
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
620
600
Arrows indicate ITD = 0 conditions for each binaural tone.
0
π
0
π
0
π
0
RELATIVE INTERAURAL TIME DELAY, ITD, AS PHASE OF 600 cps TONE, (radians)

FIGURE 5.2

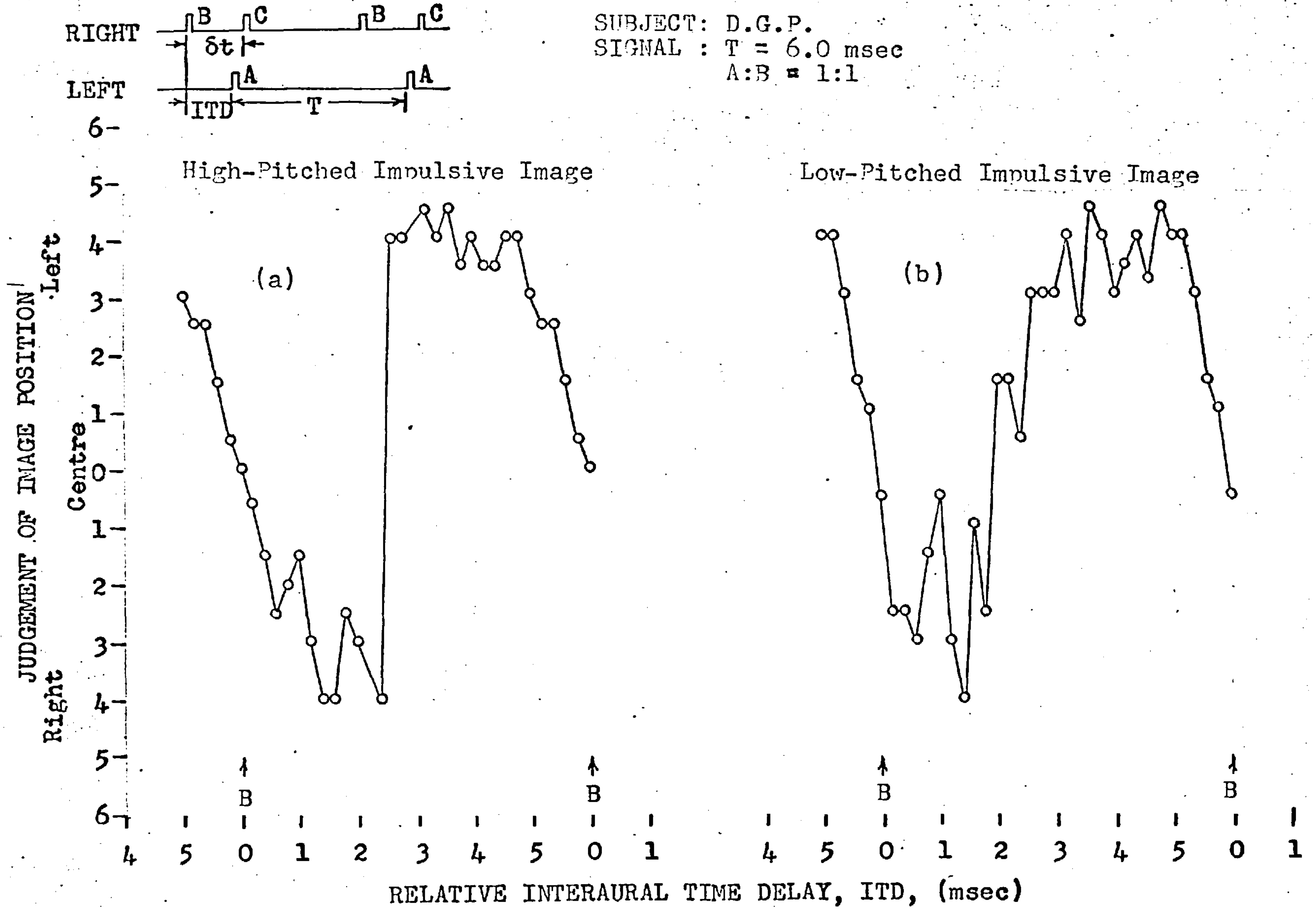
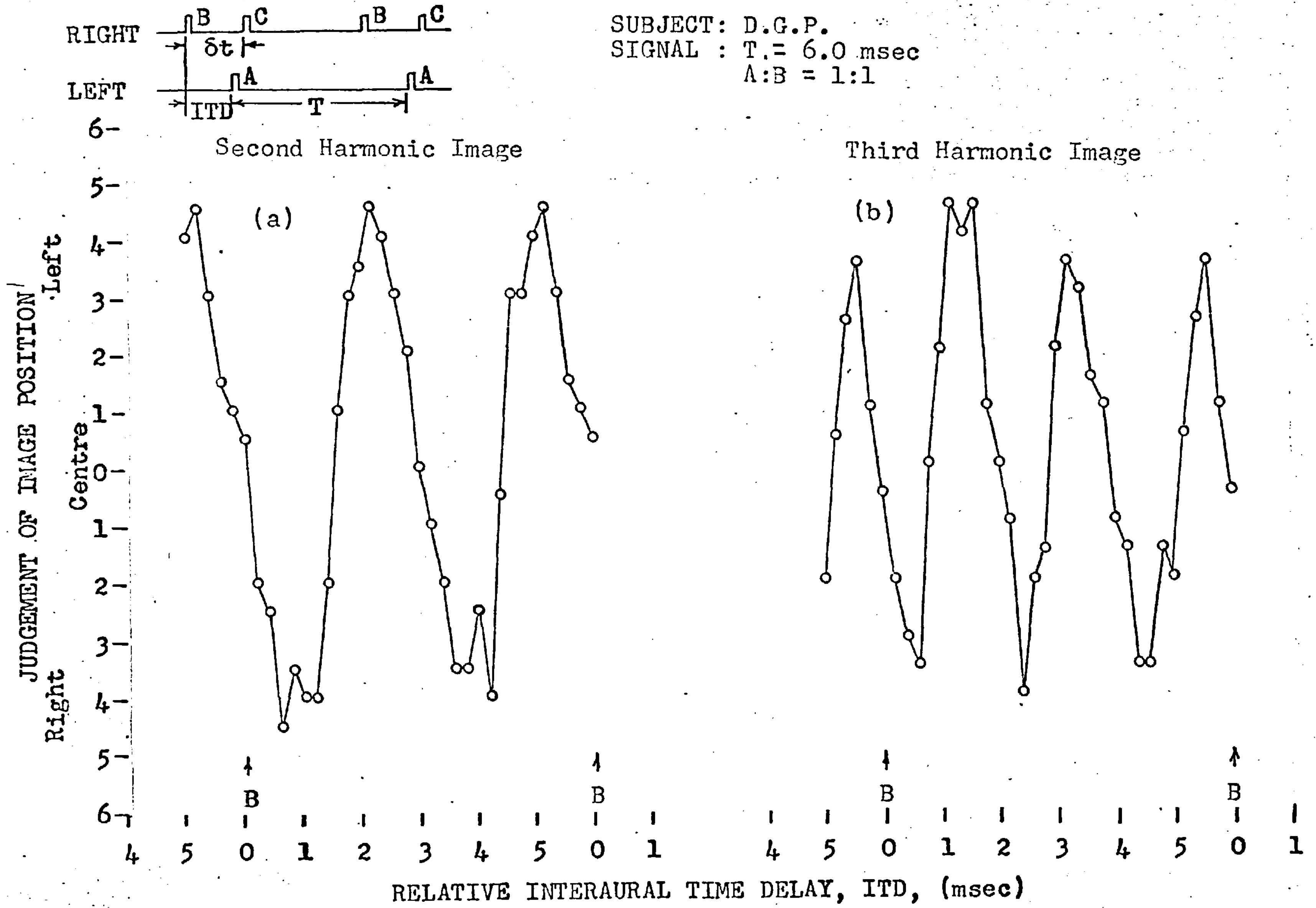
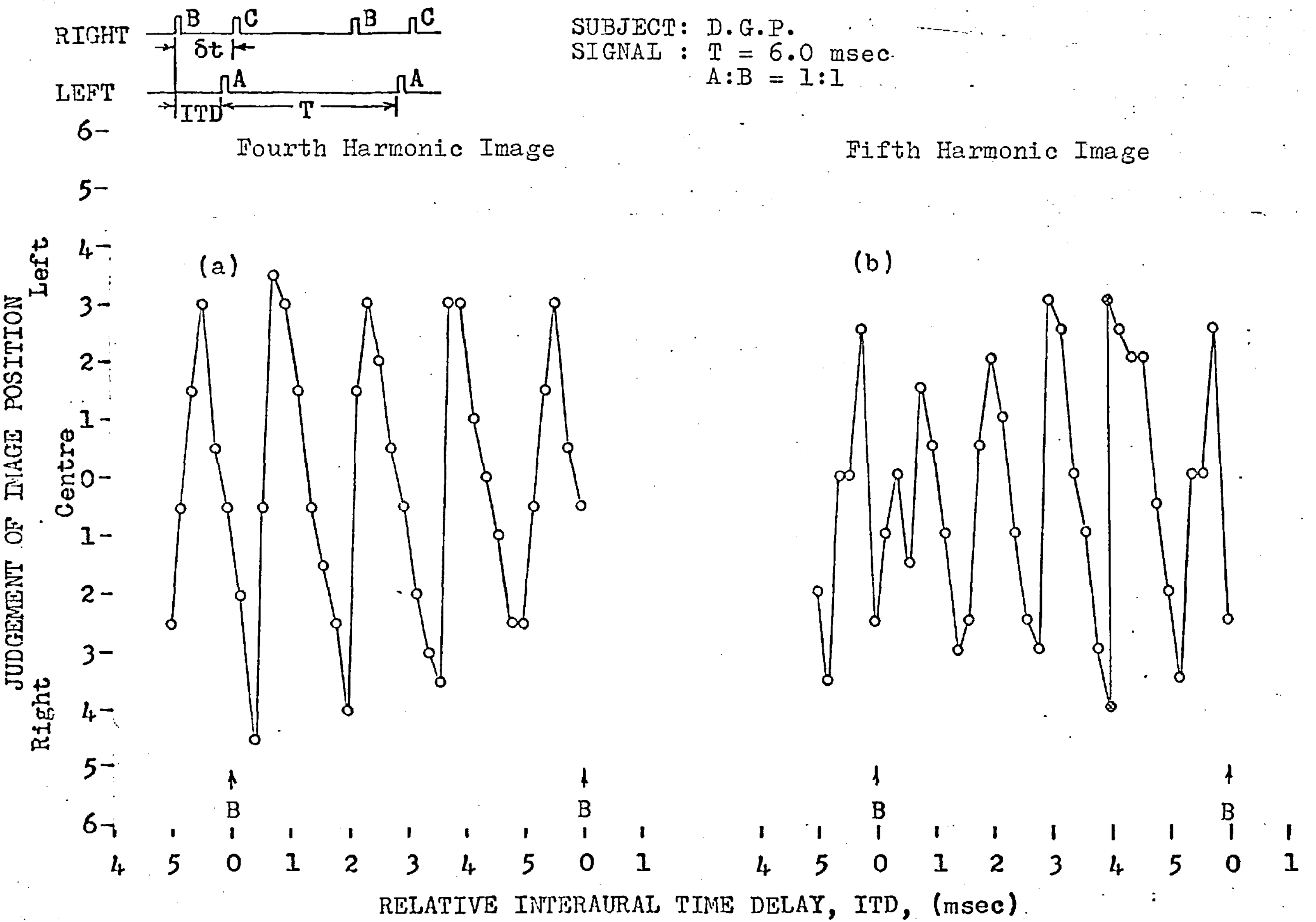

RIGHT
B
C
B
C
δt
LEFT
A
A
ITD
T
SUBJECT: D.G.P.
SIGNAL : T = 6.0 msec
A:B = 1:1
High-Pitched Impulsive Image
Low-Pitched Impulsive Image
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
B
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.3



SUBJECT: D.G.P.
SIGNAL : T.= 6.0 msec
A:B = 1:1

FIGURE 5.4


RIGHT
LEFT
B
C
B
C
δt
A
A
ITD
T
SUBJECT: D.G.P.
SIGNAL : T = 6.0 msec
A:B = 1:1
Fourth Harmonic Image
Fifth Harmonic Image
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
B
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

<u>FIGURE 5.5</u>

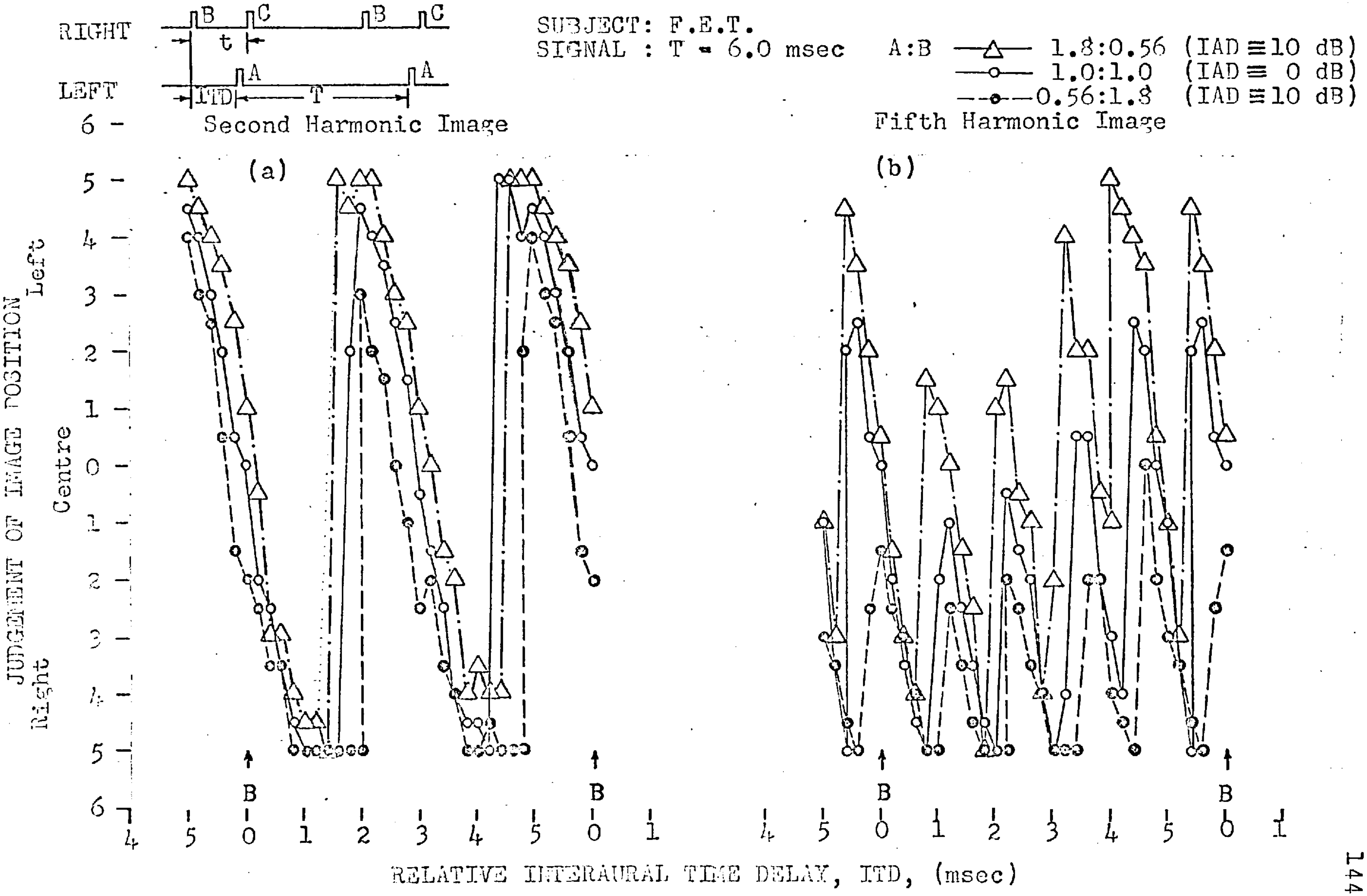

RIGHT
B
C
B
C
t
LEFT
A
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec
A:B
1.8:0.56 (IAD ≡ 10 dB)
1.0:1.0 (IAD ≡ 0 dB)
0.56:1.8 (IAD ≡ 10 dB)
Second Harmonic Image
Fifth Harmonic Image
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
B
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.6

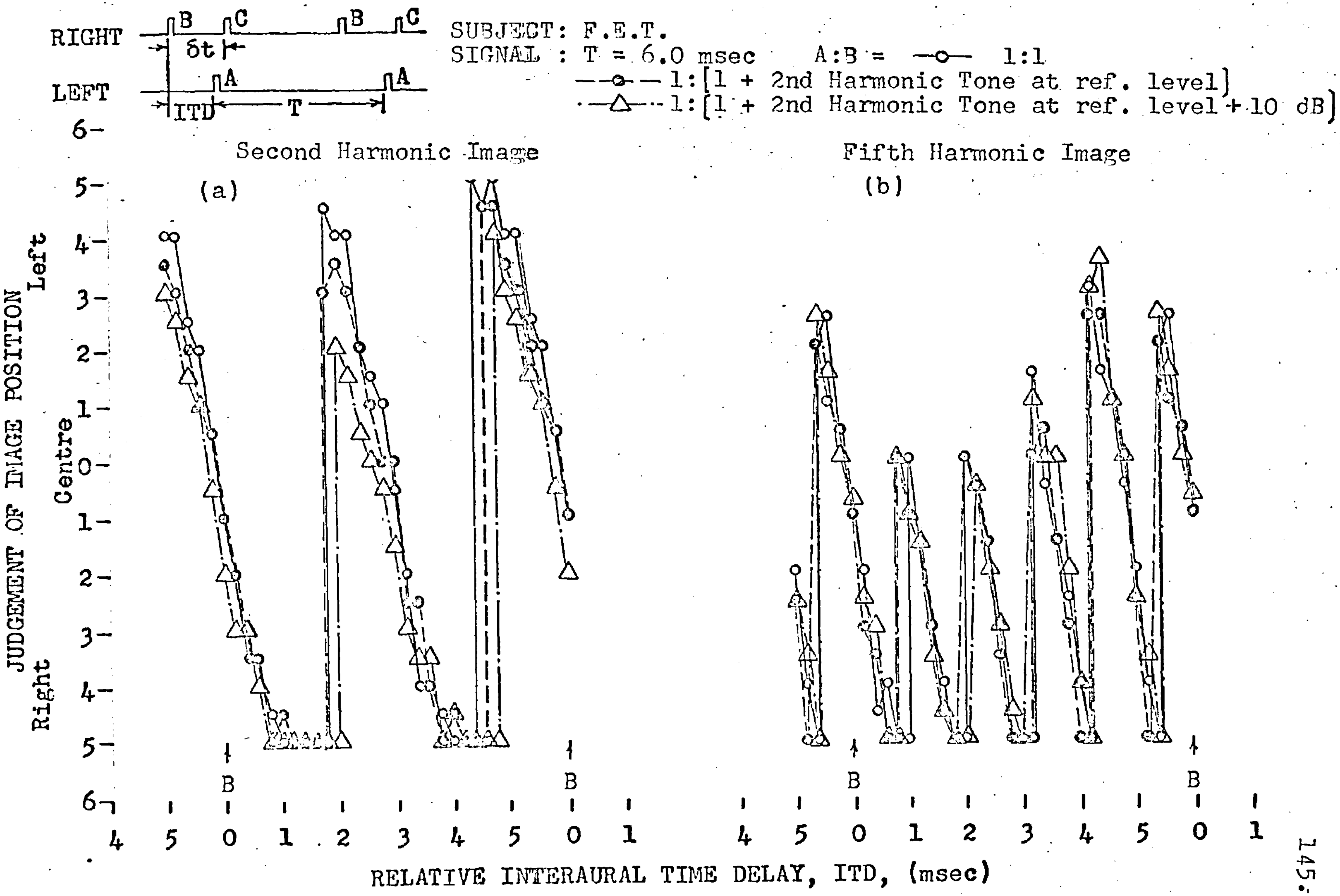

RIGHT
B
C
δt
LEFT
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec A:B = —o— 1:1
—o— 1:[1 + 2nd Harmonic Tone at ref. level]
—△— 1:[1 + 2nd Harmonic Tone at ref. level + 10 dB]
Second Harmonic Image
(a)
Fifth Harmonic Image
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
B
B
4 5 0 1 2 3 4 5 0 1
4 5 0 1 2 3 4 5 0 1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.7

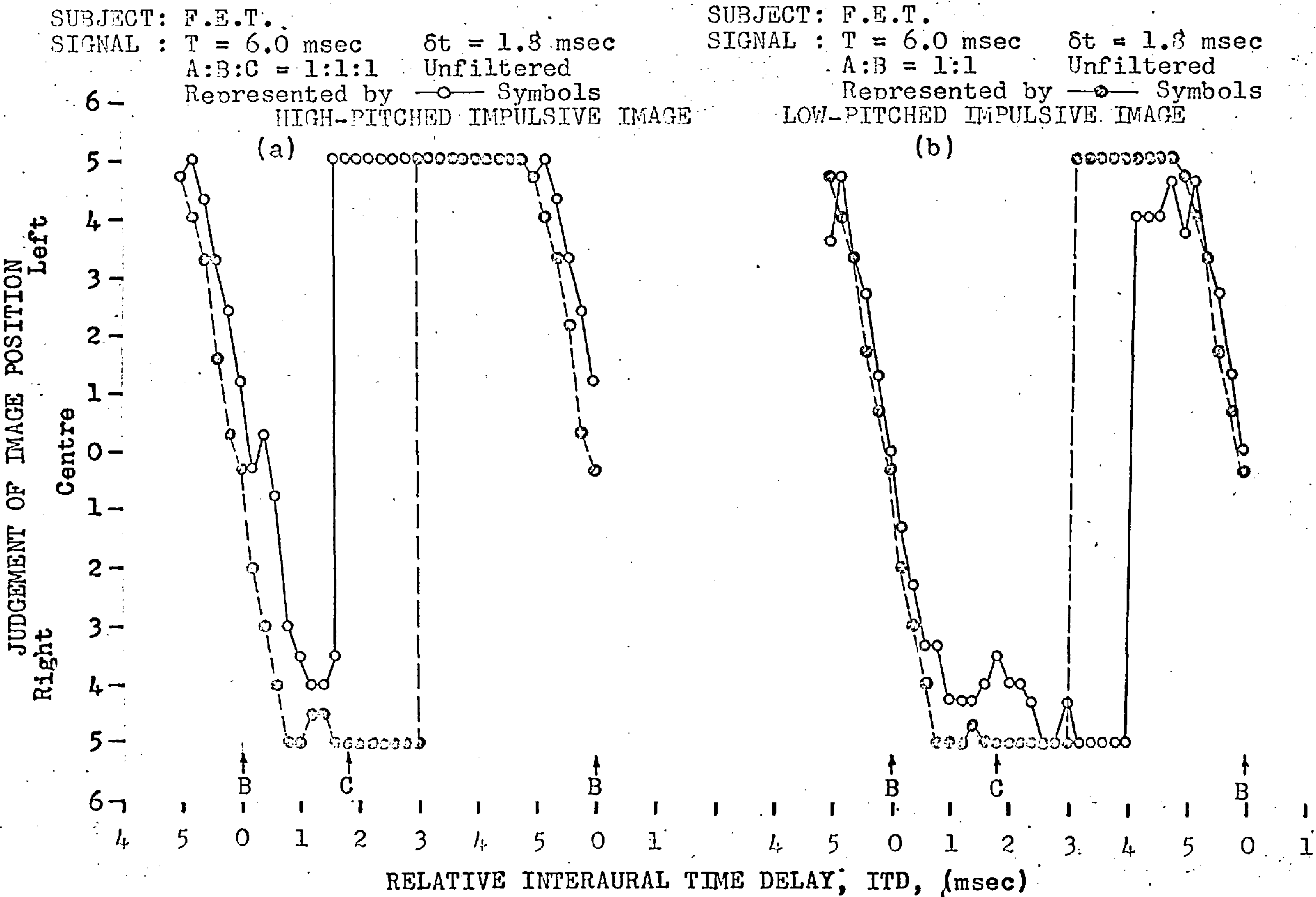

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C = 1:1:1 Unfiltered
Represented by —o— Symbols
HIGH-PITCHED IMPULSIVE IMAGE
(a)
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B = 1:1 Unfiltered
Represented by —ø— Symbols
LOW-PITCHED IMPULSIVE IMAGE
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.8

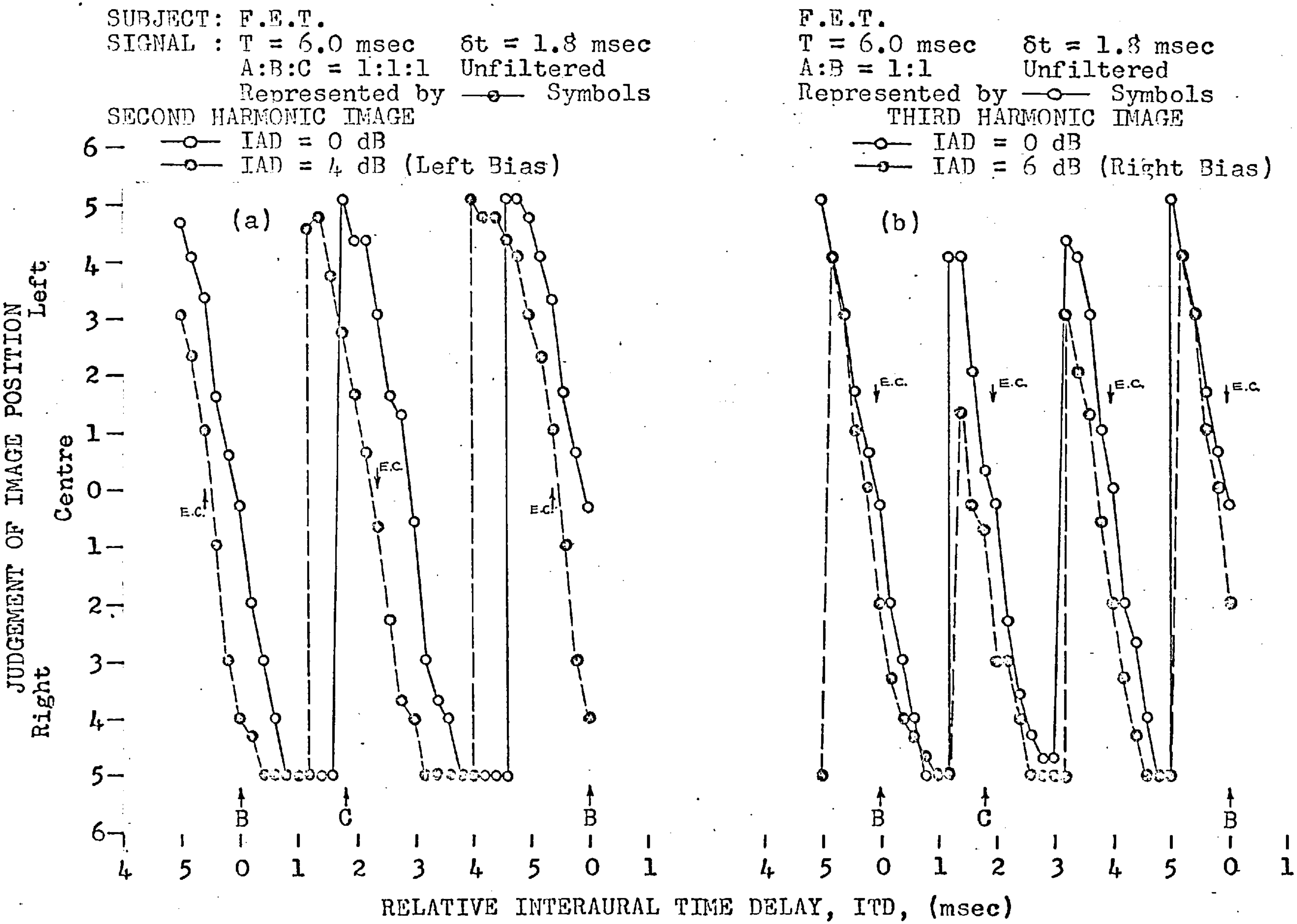

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C = 1:1:1 Unfiltered
Represented by —o— Symbols
SECOND HARMONIC IMAGE
—o— IAD = 0 dB
—o— IAD = 4 dB (Left Bias)
F.E.T.
T = 6.0 msec δt = 1.8 msec
A:B = 1:1 Unfiltered
Represented by —o— Symbols
THIRD HARMONIC IMAGE
—o— IAD = 0 dB
—o— IAD = 6 dB (Right Bias)
(a)
(b)
E.C.
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.9

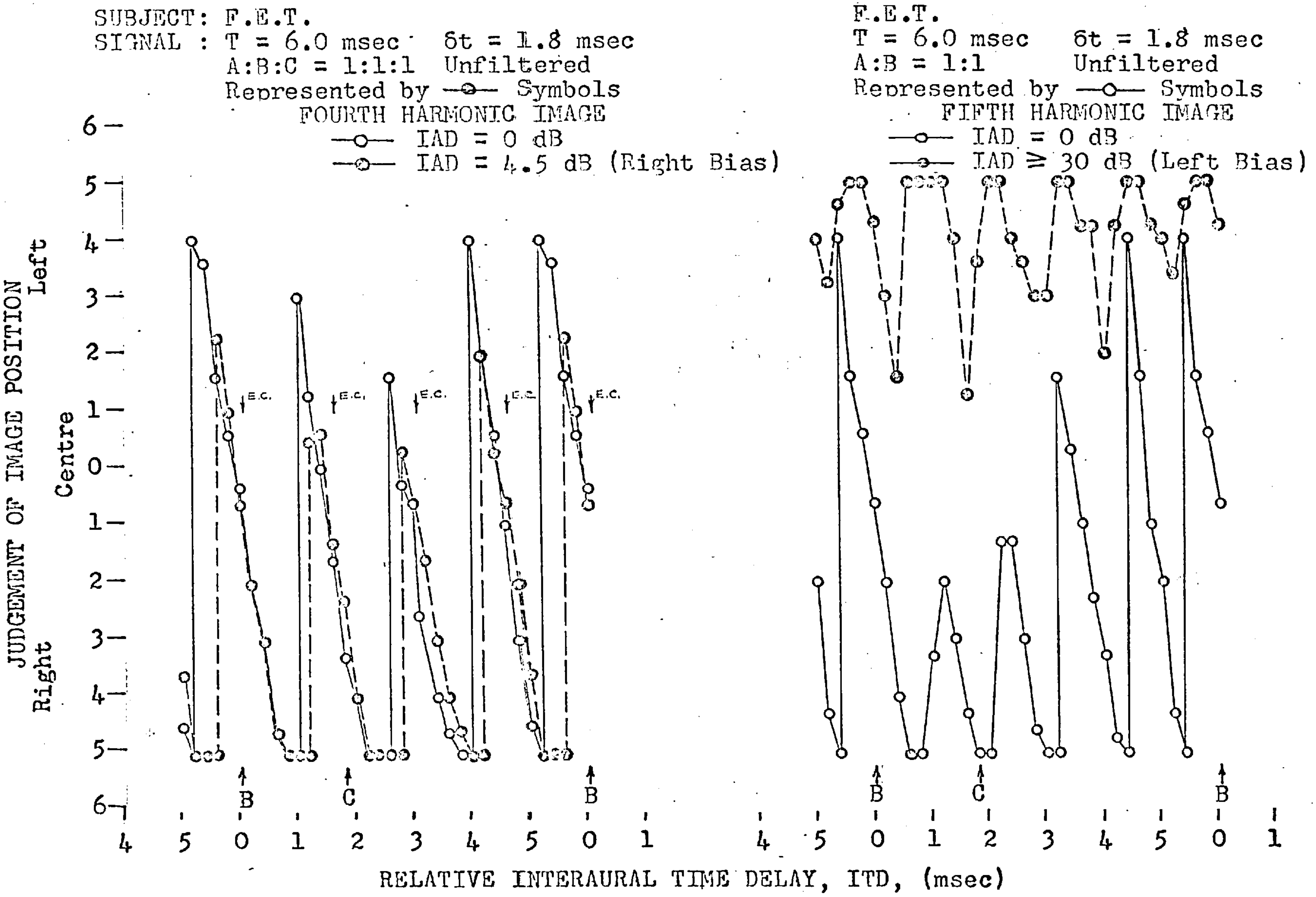
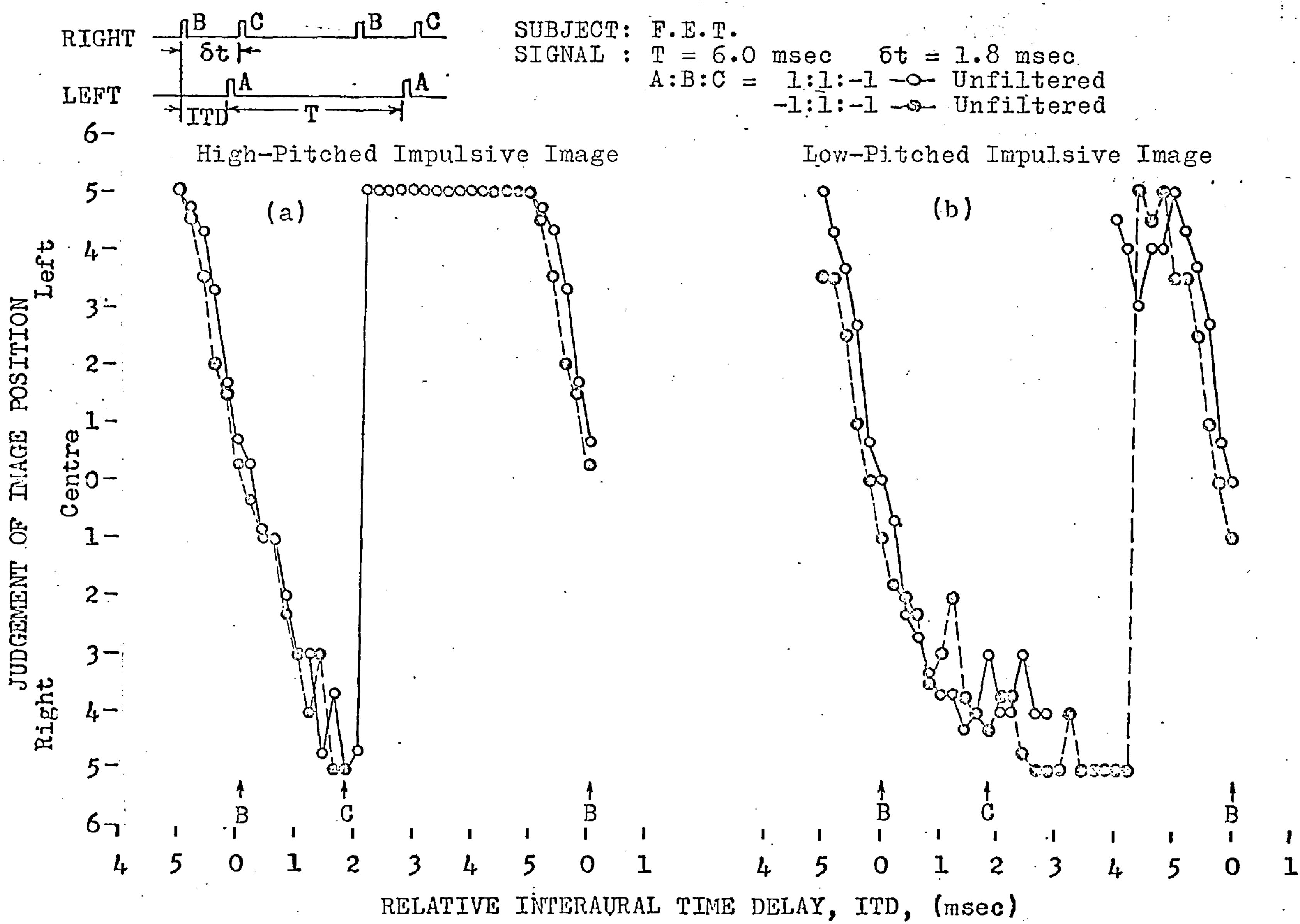

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C = 1:1:1 Unfiltered
Represented by —ʘ— Symbols
FOURTH HARMONIC IMAGE
—○— IAD = 0 dB
—ʘ— IAD = 4.5 dB (Right Bias)
F.E.T.
T = 6.0 msec δt = 1.8 msec
A:B = 1:1 Unfiltered
Represented by —○— Symbols
FIFTH HARMONIC IMAGE
—○— IAD = 0 dB
—ʘ— IAD ≥ 30 dB (Left Bias)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
E.C.
E.C.
E.C.
E.C.
E.C.
B
C
B
B
C
B
4 5 0 1 2 3 4 5 0 1
4 5 0 1 2 3 4 5 0 1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.10



RIGHT
B
C
B
C
δt
LEFT
A
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C = 1:1:-1 —o— Unfiltered
-1:1:-1 —ɵ— Unfiltered
High-Pitched Impulsive Image
Low-Pitched Impulsive Image
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.11

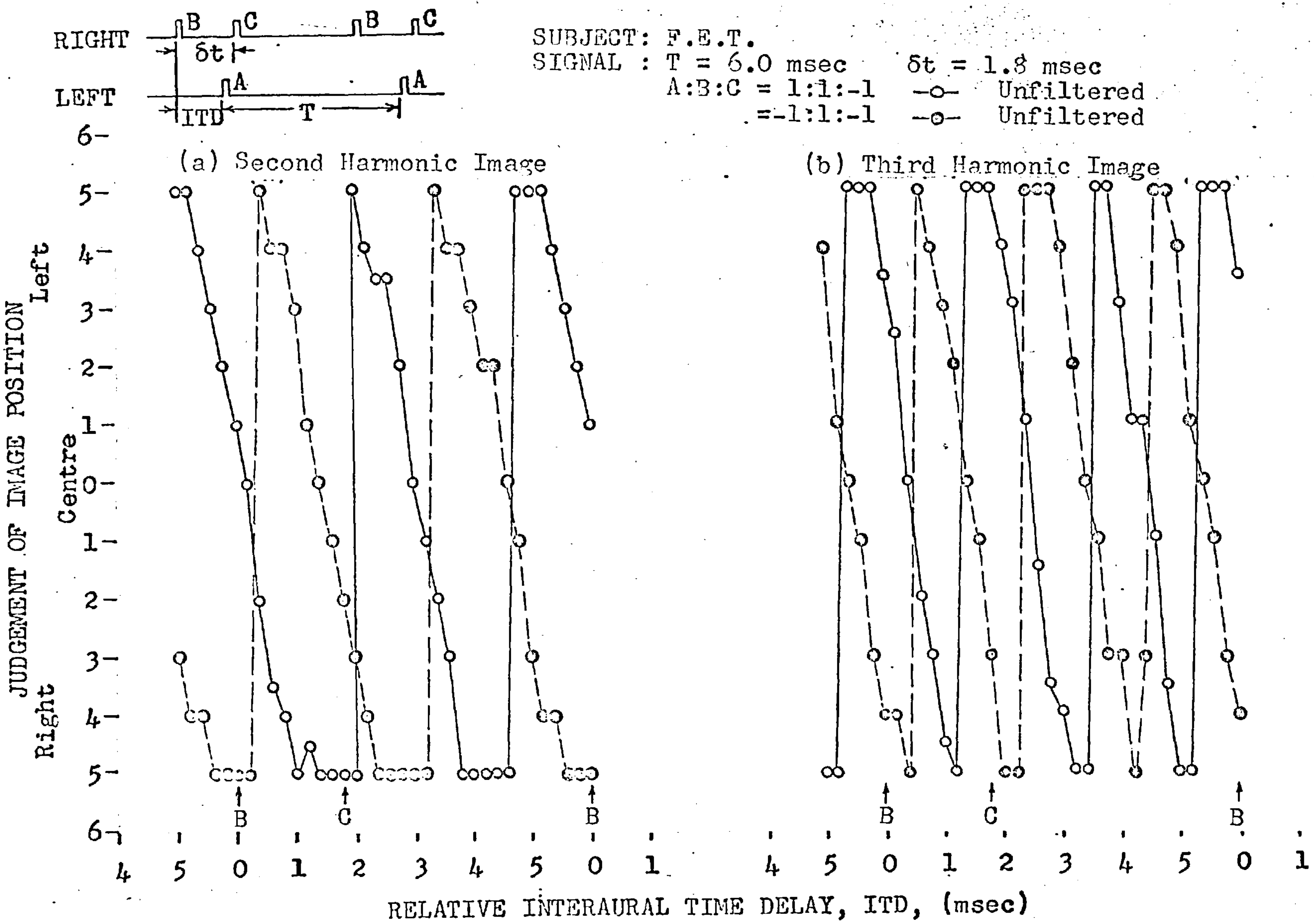

RIGHT
B
C
B
C
δt
LEFT
A
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C = 1:1:-1 —o— Unfiltered
=-1:1:-1 --o-- Unfiltered
(a) Second Harmonic Image
(b) Third Harmonic Image
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.12

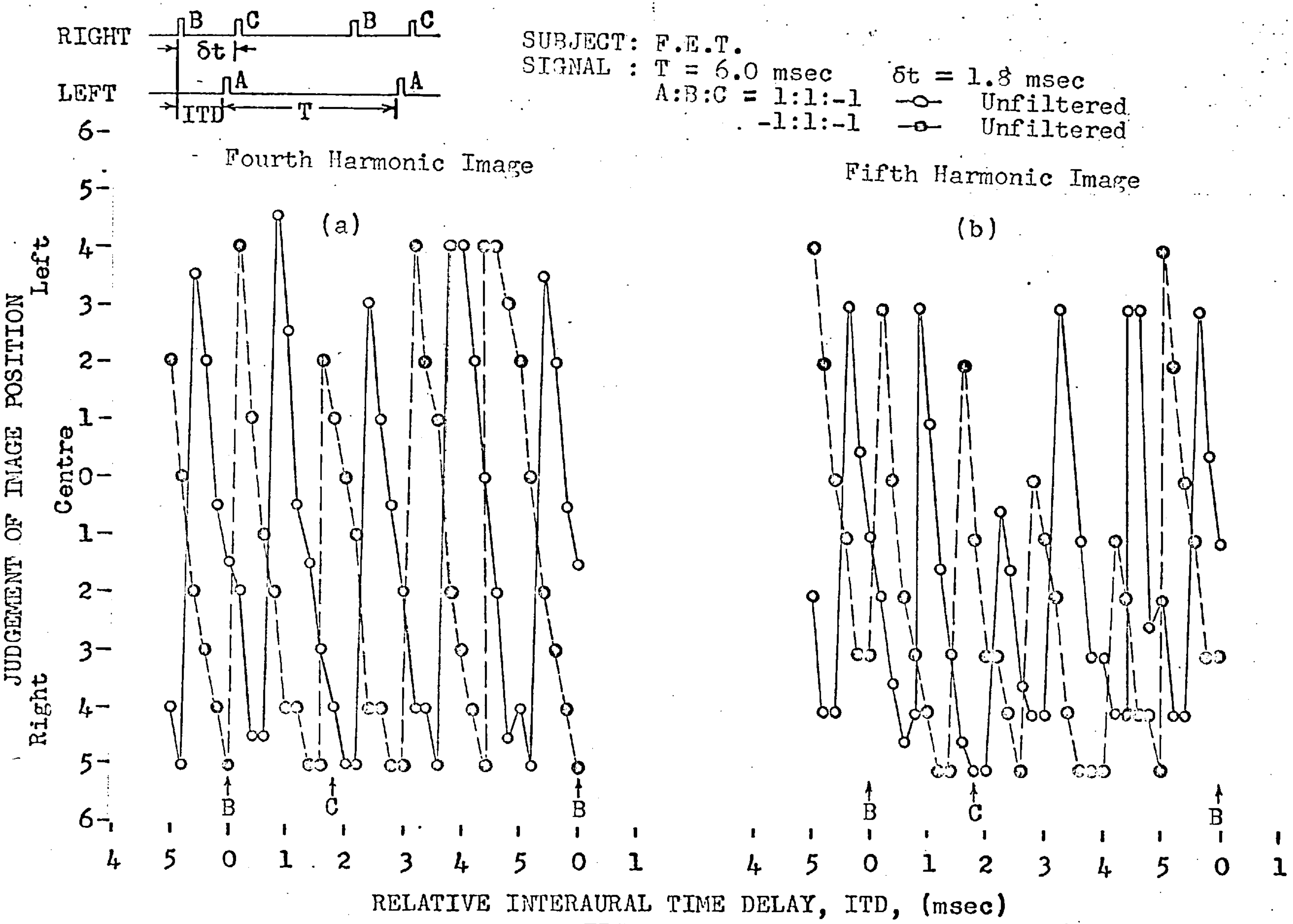
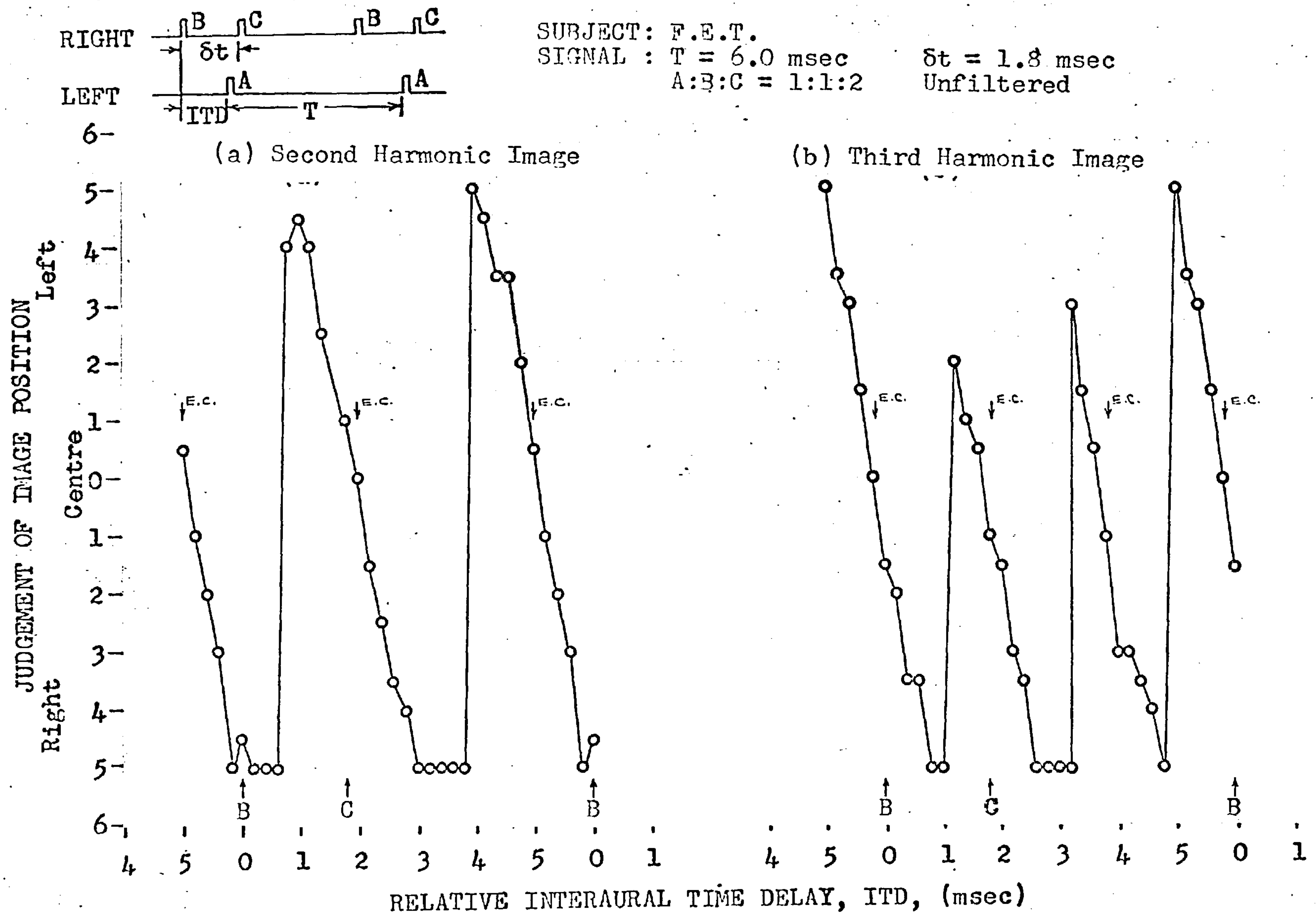

RIGHT
B
C
B
C
δt
LEFT
A
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C = 1:1:-1 Unfiltered
-1:1:-1 Unfiltered
Fourth Harmonic Image
Fifth Harmonic Image
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.13



RIGHT
B
C
δt
B
C
LEFT
A
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec   δt = 1.8 msec
A:B:C = 1:1:2   Unfiltered
(a) Second Harmonic Image
(b) Third Harmonic Image
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
E.C.
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.14

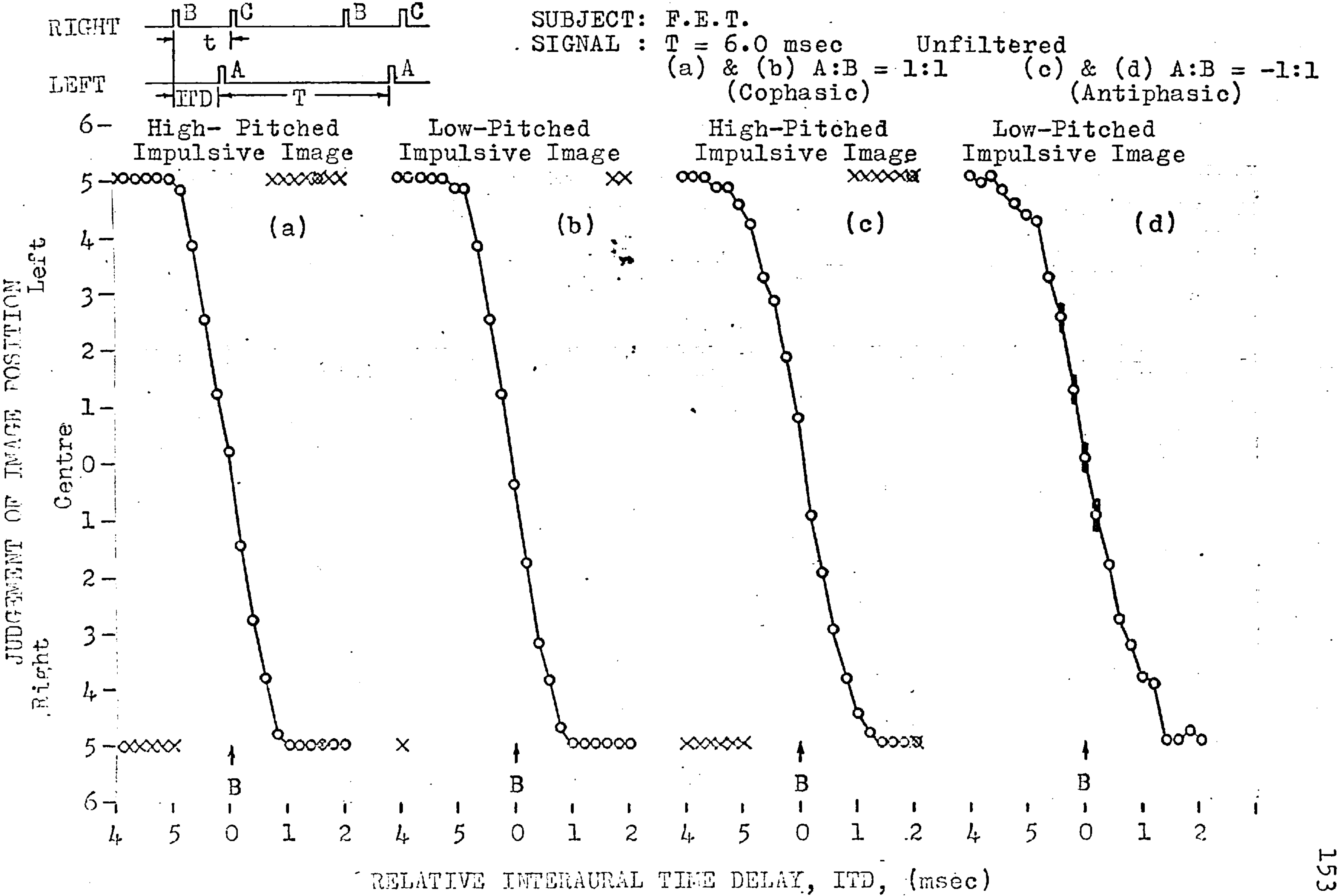

RIGHT
LEFT
B
C
B
C
t
A
A
ITD
T
SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec Unfiltered
(a) & (b) A:B = 1:1
(Cophasic)
(c) & (d) A:B = -1:1
(Antiphasic)
High- Pitched Impulsive Image
Low-Pitched Impulsive Image
High-Pitched Impulsive Image
Low-Pitched Impulsive Image
(a)
(b)
(c)
(d)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
B
B
4
5
0
1
2
4
5
0
1
2
4
5
0
1
.2
4
5
0
1
2
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 5.15

CHAPTER 6

A STUDY OF THE IMPULSIVE IMAGES

6.1.0 Introduction

It has been shown in the previous chapter that with repetitive transient signals listeners were able to identify and lateralize images of impulsive character and images of tonal character. The images of tonal character have been well described as arising from the low-frequency harmonic components of the binaural repetitive pulse patterns. However, there is convincing evidence to suggest that the images of impulsive character are the product of some substantially independent operation on the acoustic signal; in fact, the trajectories of these images indicate that they may be related to certain general features of the acoustic waveform. A complication arising from these earlier experiments was that the high-pitched and the low-pitched impulsive images were seen to exhibit somewhat different judgement trajectories, although both were apparently related to general waveform features. This fact, while attesting the independent existence of these images implied as well that they also arose from somewhat independent sources.

The purpose of this sequence of experiments was to explore by various methods the extent of the apparent waveform dependence of impulsive images and to attempt to elucidate the mechanisms which underly the perceptions of the two impulsive images.

6.2.0 Experimental Technique

In addition to the now standard experimental technique a few innovations have been introduced which allow for more powerful strategies to be applied, viz., band-pass filtering of the acoustic signal, and masking of the acoustic signal with random noise. These schemes are well known and have been usefully applied in a wide variety of experiments by other workers.

Bandpass filtering of the signal is a relatively straightforward procedure. The object is simply to remove from the signal all spectral components outside a selected band. Those frequencies within the 'pass-band' would, ideally, be transmitted without modification. In practice it is impossible to achieve the ideal, but by use of certain types of filters it is possible to closely approximate the required transmission characteristic. The main limitations of realizable filters are that the rate of attenuation between the pass-band and the stop-band is gradual and not abrupt as would be ideal, and that the spectral components of the signal undergo substantial phase shifts. In most experiments, however, these matters were of little importance as the filters were employed only to modify the acoustic waveforms as measured using an artificial ear.

The masking of signals by bands of noise is a rather different process and is, in many respects, the more useful of the two procedures. In order to provide a background for the description of the masking process, it is necessary to briefly consider a few of the underlying physiological principles.

Repetitive low-frequency signals have been observed normally to give rise to roughly synchronized volleys of neural impulses in the auditory nerve, although individual nerve fibers of the auditory nerve may show large variations in latency. There is reason to believe that "the size of the whole-nerve response is determined primarily by the probability of the auditory nerve fibers responding to each cycle of the sound wave" (Tasaki, 1954).

Individual nerve fibers exhibit the normal properties of firing thresholds and refractory periods, and it is apparently because of these refractory periods that the phenomenon of masking is operative. If random noise is added to a low-frequency repetitive acoustic signal such that the two signals give rise to overlapping regions of activity on the basilar membrane, the synchronized action potentials arising in response to the repetitive signal would be diminished by the presence of the noise. Stimulation of the nerve fibers due to the random noise would result in fewer non-refractory nerve fibers being available for synchronized firing, (Davis, 1959). A brief historical survey of the phenomenon of masking is provided by Small and Campbell (1961).

Recent studies of whole-nerve action potentials arising from acoustic transients have convincingly demonstrated that band-limited noise as a masking signal "acts primarily by eliminating portions of the normal response at times appropriate to the frequency characteristics and level of the noise. When the frequency characteristics of the noise are held constant and level is changed, higher levels remove more AP (Action Potential), but they do not appear to change the timing of the masked AP that remains. The effect is purely subtractive and is not accompanied by any other interaction between noise and transient" (Teas, Eldredge, and Davis, 1962).

An important difficulty with the procedure is that it is not possible to determine the exact portion of the basilar membrane which is stimulated by either the signal or the masking noise. This is particularly true for narrow bands of noise and signals of restricted spectrum.

The dynamic properties of the cochlear partition have been the subject of a number of studies and it is not intended to review them in detail here; some important examples are: von Békésy (1960), Tonndorf (1962), Flanagan (1960, 1962), Stevens (1951) and references cited therein.

In brief, the motion of the basilar membrane in response to a single sinusoid of displacement at the stapes covers a region of the membrane which is a function of the frequency and amplitude of the signal. All signals within the audible range are believed to cause some activity in the basal portion of the membrane (nearest to the oval window). The displacement is thought to become progressively greater out to some more distal point which is characteristic of the signal frequency; beyond this maximum the motion is rapidly diminished.

Random noise, transient signals and more complex tonal signals give rise to still more complicated response patterns, but, in general, it is believed that they induce membrane activity which broadly reflects the spectrum of the signal. High-frequency signals produce activity only in the basal portion of the cochlea; however, low-frequency signals may give rise to activity which extends from the oval window to some more distal point determined primarily by the spectral component of lowest frequency contained in the signal. In the latter case, however, the view is commonly held that the most significant neural activity arises from the maximally responding, apical portion of the organ of Corti (Flanagan, 1962).

Unfortunately, detailed information directly relating basilar membrane activity to non-sinusoidal signals is lacking. Von Békésy has provided data relating to pure tones, but information pertaining to other signals has been mainly acquired on physical or computational models.

of the basilar membrane. See, for example, Tonndorf (1962), Flanagan (1960, 1962) , and references cited therein. Certain information, however, may be inferred from observation of the cochlear microphonic waveform as measured at various points along the cochlea. It has been shown (von Békésy, 1951, see also Davis, et al., 1952) that the magnitude of the cochlear microphonic is, up to a limit, closely proportional to the displacement of the cochlear partition. From measurements of this type, in fact, von Békésy's direct observations of mechanical 'filtering' have been fundamentally confirmed (Tasaki, et al., 1952).

Teas, Eldredge and Davis (1962) have measured both cochlear microphonics and whole-nerve action potentials in the various turns of the cochlea, and have by use of selective masking procedures derived the action potentials which are presumed to arise in regions of the membrane corresponding to various frequency bands. To avoid misleading results which might possibly have arisen due to low-frequency bands of noise stimulating the same regions of the basilar membrane as high-frequency bands, the authors used high-pass filtered noise throughout, varying only the cut-off frequency and thus, presumably, the distal limit of random membrane activity.

The sum of the derived action potentials so closely approximated the whole action potential that these authors were led to conclude that the masking noise simply subtracts from the neural population which responds to the transient stimulus, and that this subtraction is effected over a region of the basilar membrane which corresponds to the spectral extent of the masking noise.

Thus it was considered useful to apply this technique of masking to the binaural transient signals employed here with a view to determining the spectral components and/or the regions of the cochlea which contribute to the various impulsive images which have been reported by listeners in these experiments. It was also of some importance to relate these reported images, as far as possible, to the current views of neurophysiological activity associated with relevant basilar membrane displacement. The use of filtered masking noise to diminish the effects of selected bands of signal spectra has the apparent advantage over signal filtering, of avoiding problems of signal contamination due to the filter characteristics.

The level at which the masking noise was presented was, unless otherwise stated, established by first presenting wide-band noise at a level just sufficient to render inaudible the wideband transient signal which was presented at the now standard experimental level of about 30 dB SL. The level thus determined was maintained and the noise spectrum then manipulated as desired by adjustment of the variable bandpass filter. Where the noise was applied binaurally, two uncorrelated noise sources were used, one for each channel, and each was independently filtered.

Listeners used in this series of experiments were those with considerable experience in the previous experiments and who showed sufficient aptitude and interest to work hard at these very difficult and tiring experiments. Results that are shown were taken with one listener to enable ready comparison, however, each important experimental result was verified by using at least three listeners. In each case results were very similar. Primary and secondary judgements of split-image situations have been indicated in most graphs. The primary (dominant image) judgements were indicated as open circles and secondary judgements by crosses.

6.3.0 Apparatus

The apparatus described by Fig. 4.1 was modified slightly to accommodate variable bandpass filters and noise generators for the purpose of filtering and selective noise masking of the acoustic signal.

Experiments in which the bandwidth of the pulse signal was restricted, utilized variable bandpass filters (Krohn-Hite 315A and Krohn-Hite 310 AB, different models but having identical filter characteristics; rate of attenuation = 24 dB/octave) inserted in each signal path prior to the summing amplifiers. All impedences were appropriately matched.

Experiments utilizing masking noise were arranged as follows: outputs of two random noise generators (Dawe,419BR and General Radio,1390B) were individually passed through variable bandpass filters and applied to inputs of the left-and right-channel summing amplifiers.

6.4.0 Experiments

6.4.1 Effects of Signal Filtering: Exploratory Experiments

Listeners, when asked to describe the character of the impulsive images, were agreed on the basic points that one had a low-pitched character and the other had a high-pitched character; although descriptions also ranged to such terms as "clicks","hums," "buzzes", etc. The first experiments to be carried out were intended to test whether these subjective impressions could be supported by experiments involving spectral manipulations of the signals.

A few casual experiments provided useful guiding information in this respect. Listeners were asked on different occasions to attend to either the low or the high-pitched impulsive image (at fixed ITD, say,=0) and to adjust the cutoff frequency of the variable band-pass filter (set to either low-pass or high-pass mode) until the character of the image was noticeably altered. In all cases listeners reported that the character of the low-pitched impulsive image was not apparently altered by signal band-limiting down to about 1000-2000 cps low-pass, nor was the character of the high-pitched impulsive image apparently modified by band-limiting up to about 1000-2000 cps high-pass. Using this evidence as a guide, more formal experiments were arranged.

Two such experiments involved either high -or low-pass filtering (cut-off frequency = 1500 cps) of the binaural signals, in this case repetitive single versus double-pulse trains, thus presumably eliminating the low-or high-pitched impulsive images respectively. Under each of these conditions listeners reported hearing only one image and when this was lateralized according to the normal experimental procedure it was found that the judged trajectories for the remaining low-or high-pitched impulsive images closely resembled those obtained for the corresponding images in earlier experiments where these were simultaneously perceived but independently lateralized (e.g., Fig. 5.8).

A more rigorous experiment was then arranged in which both high-and low-pitched impulsive images were simultaneously perceived and individually lateralized. In this experiment the members of the pulse-pair, B and C, were individually band-limited; pulse A was unfiltered and either pulse B or pulse C was low-pass filtered (cutoff frequency = 1500 cps) with the remaining member of the pair unfiltered. Figs. 6.1 and 6.2 illustrate judged image trajectories for the low-and high-pitched impulsive images when the B and the C pulses respectively were low-pass filtered. For comparison Fig. 6.3(a) depicts the result when both B and C pulses were low-pass filtered. The high-pitched impulsive image was reported continuously at position 5 left in the latter experiment.

An important feature of these results was that the judged trajectories of the high-pitched impulsive image were very clearly dominated in each of the experiments by the unfiltered acoustic transients. Reference to Fig. 5.8 which illustrates the results of a similar experiment but where all pulses were unfiltered, suggests that the filtering of the C pulses introduced no important change in the high-pitched image trajectory; however, filtering of the B pulses altered the judgement pattern substantially, and in an important fashion.

Trajectories of the low-pitched impulsive image also showed some interesting changes, however, these changes seemed to be of a somewhat different kind. The manipulation illustrated in Fig. 6.2 apparently did not seriously modify the judged trajectories of the low-pitched image (compare with Fig. 5.8). The judgement curve illustrated in Fig. 6.1, however, exhibited some modifications which are not easily accounted for, viz., the unusually well-defined tracking of both of the S_1 and S_2 images, and the unexpected ITD values at which these trajectories were seen to cross the centre axis. In a sense, the results of Fig. 6.3(a) could be interpreted as also reflecting a rather well-defined tracking of the S_2 image, in fact, to an almost equal extent as the S_1 image. Note the displacement of the centre crossing.

These results seemed to support subjective data suggesting that the high-pitched impulsive image was associated with spectral contributions above 1500 cps and that the low-pitched impulsive image was associated with spectral contributions below about 1500 cps. Some modification of the low-pitched image trajectories was observed; the immediately obvious possibility was that this occurred as a consequence of the technique of signal filtering and the modification of acoustic waveforms which resulted.

To clarify this last matter, a set of experiments was arranged which used a different acoustic waveform in each channel, but in a less complex signal configuration, in this case binaural single-pulse trains.

Figs. 6.3(b) and 6.4 illustrate results of three such experiments, along with sketches of the relevant acoustic waveforms as measured on the artificial ear (shown to the same time scale as the ITD). Judgement trajectories for the low-pitched impulsive image are seen to have been progressively displaced along the ITD axis as the right-ear pulse was low-pass filtered with cutoff frequencies of 15000 cps , 1500 cps and 600 cps. Relative to the 15000 cps low-pass (i.e., wideband) case, the displacements were about 300 μsec for the 1500 cps low-pass signal, and about 800 μsec for the 600 cps low-pass signal. Reference to the measured acoustic waveforms and in particular to the interaural time relations between the first peaks of left-and right-ear signals in the relevant experiments, indicates close agreement with the observed ITD values for judged image centre crossings. This would appear to further confirm and extend the waveform-dependence of the impulsive images hypothesized earlier, and to provide what seemed to be a possible explanation for the delayed centre crossing of the low-pitched S_1 image trajectory in the experiments of Figs. 6.1(b) and 6.3(a).

In the experiments of Figs. 6.3(b) and 6.4, the perceived characters of the sound images were altered drastically by the filtering operations, yet they were readily and reproducibly lateralized by a number of listeners. As the high frequency components of the transient signal in the right ear were removed by filtering, listeners reported an impulsive image of high-pitched character, fixed at the left extreme position. The pitch-like character of this image was perceived to alter with filter cutoff frequency, as applied to the right-ear signal. An interesting observation was that in the 600 cps low-pass condition, the character of the normally-dominant low-pitched impulsive image was altered and that a 'residue' image of low-pitched character was generally perceived at the left lateral extreme along with a high-pitched impulsive image. Each of these images remained fixed in position as ITD was altered. It was concluded that these images arose as a consequence of the unilateral filtering operation and that the positionally fixed images were associated with spectral components in the stop-band of the filter. It seemed that neural activity associated with these spectral bands either did not interact (in which case the static images would be monaural images), or in the interaction were subject to an effective IAD of such magnitude as to cause complete lateral displacement of the resulting image. The term 'residue' seemed an apt description for the phenomenon.

It has been stated above that only two impulsive images had been perceived to coexist. Indeed, listeners had not normally reported more than this number in experiments using wideband signals. Here, however, was a case where listeners reported hearing the familiar high-pitched impulsive image fixed at one lateral extreme;

and <u>two</u> images of low-pitched impulsive character; one apparently fixed in position at the left extreme and another lateralized with ITD in the normal manner. The characters of the low-pitched images were perceived to alter when the low-pass cutoff frequency was varied about 600 cps; the fixed image was subjectively associated with frequencies above the cutoff frequency, and the movable image with frequencies below the cutoff frequency. A similar phenomenon was reported by the listeners who participated in the casual experiments which opened this series.

This finding suggested that the low-and high-pitched impulsive images commonly identified may be only two of many such images, or may represent combinations of multiple impulsive images. In any event there was a case for exploring such a possibility.

An observation which is of interest in this matter came out of experiments utilizing repetitive single- versus double-pulse trains. If a listener was given control over the ITD parameter, the low-pitched impulsive image could be tracked along the hypothesized S_2 image trajectory, crossing centre at closely ITD = δt, in conditions where the results of a formal image lateralization experiment suggested that this image was only of lesser influence, for example, in the transitional region between S_1 and S_2 trajectories at small $\delta t/T$ values. When, in such situations, the signals were removed and then, after a short interval, re-applied, the listener often reported an image position which corresponded to that which may have been judged in the experimental context. In short, a listener may, by himself adjusting the ITD, be able to track a low-pitched image which apparently follows the S_2 image path in a situation where normal experimental results suggest this is not to be expected. Momentary interruption of the signal, however, may cause the listener to apparently lose track of this image, and report a positional judgement which is characteristic of that expected within the formal experimental context. Occasionally, by dint of considerable effort, some listeners appeared to be able to again seek out this rather elusive image and report it as being in its novel position.

These results led to the introduction of a new experimental strategy, the intention of which was to direct the listener's attention to the low-pitched impulsive image in an experiment where the high-pitched image also was normally perceived. This strategy took the form of a series of experiments arranged so as to at first allow the listener to lateralize the low-pitched image in isolation and thereafter amidst gradually increased distracting influences.

Results of the set of six experiments are illustrated in Figs. 6.5, 6.6 and 6.7. The signal configuration chosen for this series was that which had been employed on a number of previous occasions, and for which many results had been accumulated. The left-ear signal was repetitive single-transients and the right-ear signal repetitive double-transients (T = 6.0 msec, δt = 1.8 msec, IAD = 0), as in Figs. 4.7, 4.22, 4.23 4.26, 5.8, which prove useful as comparative results.

The first experiment used the binaural signals low-pass filtered with a cutoff frequency of 1500 cps. Fig. 6.5(a) illustrates the result, which is only slightly different from those observed with wide-band binaural signals.

The second step was to reduce the cutoff frequency to 600 cps. Fig. 6.5(b) illustrates the result of this manipulation, again, quite similar to Fig. 6.5(a) and to the earlier results mentioned above.

In these two experiments the listener reported little difficulty in performing the set task; however, when the right-ear signal bandwidth was increased by low-pass filtering with a cutoff frequency of 1500 cps the difficulty was increased. Fig. 6.6(a) shows the judged trajectory for the low-pitched impulsive perceived image in this experiment. There is a very definite improvement in the apparent resolution of the S_1 and S_2 images here, in spite of a steady right-extreme residual image reported by the listener.

Fig. 6.6(b) illustrates the result of an even more difficult experiment where the right-ear signal was wideband and the left-ear signal still 600 cps low-pass filtered. Again, the S_1 and S_2 image trajectories are seen to stand out very clearly, and, indeed, appear to have been lateralized in a fairly idealistic manner. Both curves of Fig. 6.6 exhibit displacements along the ITD axis, presumably due to the different acoustic waveforms in each channel.

To test the listener's performance under normal symmetrical bandwidth stimulation the experiments of Fig. 6.7 were arranged. In the experiment of Fig. 6.7(a) both signals were 1500 cps low-pass filtered, and in Fig. 6.7(b) both signals were wideband. In each case the S_1 and S_2 image trajectories were almost equally represented in the pattern of judgements. The listener reported the last experiment to be extremely difficult, the low-pitched image often proving to be quite elusive amongst the array of harmonic tonal images and the higher-pitched impulsive images. The preparatory experiments were reported to have been of considerable assistance in recognizing the image to be judged.

Certainly the results would seem to confirm this impression; comparison of the curve of Fig. 6.7(b) with any of those of earlier experiments of similar nature which were conducted in isolation established the degree of this change in performance. Within this set of experiments a comparison of Fig. 6.5(a) with Fig. 6.7(a) is sufficient evidence; the only difference in the circumstances of these experiments was the formal preparation which preceded the latter.

A number of interesting points have emerged from the results so far, namely:

(1) a listener's ability to resolve certain of the different impulsive images can apparently be enhanced by suitable training.

(2) the reliable identification of the different impulsive images which are perceived in certain experiments may prove to be a key factor in elucidating the mechanisms which underlie these images.

(3) the proposition that monaural temporal masking is the main factor which determines the degree to which the S_1 and S_2 images contribute to the pattern of impulsive image judgements, is clearly incomplete. The well-defined tracking of both S_1 and S_2 low-pitched images seen in Fig. 6.7 is rather different from comparable judgement trajectories seen in earlier experiments (e.g., Fig. 4.7).

(4) the conclusion that in these experiments there are only two impulsive images of distinctly different 'timbre' may also be in some doubt. It is possible that there are many such images, and that according to the

experimental conditions or listener training, one or more of these may influence the listener's impression sufficiently to be individually identified and lateralized.

The experiments which follow were therefore arranged with the intention of clarifying one or more of these issues.

Perhaps the most important question which arises is that which bears on the role of monaural masking in determining the relative importances of the S_1 and S_2 images in the judgement process. Results reported in Chapter 4 showed what appeared to be a systematic reduction in the relative importance of the S_2 image as δt was reduced relative to T. In the experiments of Chapter 5 where listeners independently lateralized what were reportedly low-and high-pitched impulsive images, both subjective impressions and experimental results suggested that the low-pitched impulsive image was the single dominant image which was lateralized in the experiments of Chapter 4. However, the previous argument of the relative influence of monaural masking seemed to relate only to the low-pitched impulsive image, because in the case of the high-pitched impulsive image the S_2 image did not appear to exercise any influence in the positional judgements. In short, the S_2 image would seem to have been completely masked (i.e., neural products of the C pulse effectively eliminated). The only feature of these results which negates such an apparently obvious conclusion was the earlier judgement transition when the C pulse was present (see Fig. 5.8(a)), therefore, perhaps, indicating an effect of a different kind.

The results illustrated in Figs. 6.1 and 6.2 also have bearing on this matter. In each of these experiments the spectrum below 1500 cps was left essentially intact in both channels while that above 1500 cps was selectively manipulated. The judged trajectories of the high-pitched impulsive image were shown to be dominated in each case by the contralateral pulses containing the higher spectral components. The pertinent feature was that in the experiment of Fig. 6.1 the low-pass filtered B pulse appeared not to exercise any important amount of masking over the contributions of the following wide-band C pulse to the high-pitched impulsive image.

It was clear that further experiments of a specialized kind were required. Because of the apparent differences in the effects of monaural masking on the low-and high-pitched impulsive images, the problem of monaural masking was considered in two parts; one part dealing with the low-pitched impulsive image phenomena, and another dealing with high-pitched impulsive image phenomena.

6.4.2 Effects of Signal Filtering: Low-Pitched Impulsive Image Phenomena

Judged trajectories of the low-pitched impulsive image illustrated in Fig. 6.7 indicated an absence of monaural masking effects such as were apparently observed in earlier similar experiments (e.g., Fig. 4.7). As an extension of these earlier results, Figs. 6.8 and 6.9 show judged trajectories for the high-and low-pitched impulsive images in a set of four interlaced experiments where δt was fixed at 2.4 msec and the repetition period was increased from 6 msec through the values of 10 msec and 20 msec to an upper limit of 40 msec. As has been the case in other experiments of this type, the trajectories of the high-pitched image reflected, largely, the existence of the S_1 image only; those of the low-pitched image showed as the repetition period was increased, a rapid reduction in the effects of the S_2 image. In the experiment of Fig. 6.9(a) (T = 40 msec), however, a few isolated judgements may be seen which could be interpreted as falling on or near the hypothesized S_2 image path.

If monaural masking was, indeed, the factor which resulted in the gradually reduced effects of the low-pitched impulsive S_2 image as the repetition period was increased, then it was extremely interesting that these few judgements should have appeared where and when they did. To test whether or not these unexpected judgements were fortuitous events, the experiment was repeated, but this time after a short preparatory period during which the listener attended to the image resulting from binaural low-pass filtered signals and then continued to seek out this image when the bandwidth restrictions were removed. After some practice this procedure could be casually carried out in the few minutes just prior to a formal experiment and, in fact, in later experiments was often unnecessary.

Fig. 6.10(a) illustrates results of an experiment of this type where T = 40 msec and $\delta t = 1.8$ msec. Judgement curves resulting from experiments using signals bandlimited to 600 cps low-pass and those resulting from experiments using wideband signals which were preceded by some amount of training were virtually indistinguishable, and are equally well represented by Fig. 6.10(a). It must be emphasized that the latter type of experiment was extremely difficult; three listeners were able to closely reproduce the illustrated result when the signals were low-pass filtered, but only one of these could reliably identify and lateralize the low-pitched S_2 image when the signals were wideband. Nevertheless the result was extremely important and added considerable support to the view that monaural masking is not an important factor in the mechanism which underlies low-pitched impulsive image phenomena.

Still more convincing were the curves of Fig. 6.10(b) and (c) where T = 100 msec and δt = 1.2 and 2.0 msec respectively. Here the signals were low-pass filtered with a cutoff frequency of 600 cps. Again, the low-pitched S_2 image is seen to be clearly tracked in each case.

Symbols used in plotting the results of these experiments have been chosen to reflect the order of the listener's judgements. In these experiments, however, the multiple images were of basically similar character, viz., they were impulsive; the distinguishing features were often quite subtle differences in timbre or 'pitch-like' attributes. When a number of these images were judged to occupy the same portion of a listener's auditory space it was very difficult, and often impossible, to realistically attempt a comparison of the relative importance of split-image components. Very little significance is, therefore, attached to ranking of the judgements in these experiments, although it is with no small respect that one considers the consistency with which they were reported.

6.4.3 Effects of Signal Filtering: High-Pitched Impulsive Image Phenomena

The high-pitched impulsive image has been shown to exhibit lateralization properties rather different from those of the low-pitched image when both images were simultaneously perceived and independently judged. The curves of Figs. 6.8 and 6.9 illustrate these differences quite well. It is first interesting to enquire what would result if this image were to be lateralized in isolation, as has been the low-pitched impulsive image. Secondly, it is of importance to enquire into the apparent absence of the high-pitched S_2 image.

A series of experiments was arranged with the purpose of gaining some insight into both of these problems. The signals employed were, as before, repetitive single versus double-pulse trains, on this occasion high-pass filtered with a cutoff frequency of 6000 cps.

Experiments of this type employing the signal configuration used earlier, where $T = 6$ msec and $\delta t =$ various, resulted in judged trajectories of the high-pitched impulsive image which were identical in all respects to those shown in the appropriate experiments for this image when lateralized in the presence of the low-pitched impulsive image.

If the repetition period was increased to 100 msec, say, and δt increased to, say, 8 msec, the interesting result obtained wherein both the high-pitched S_1 and S_2 image trajectories appeared in their entirety. Fig. 6.11(a) illustrates this result. In the ranges of ITD outside those in which the main image crossing occurred the listener was unable to judge with confidence the relative amplitudes of the split images at the contralateral extreme positions, hence, no judgement transition is shown.

In the following experiments the repetition period was held constant at 100 msec and δt was reduced by steps. Fig. 6.11(b) illustrates results for $\delta t = 4.5$ msec; Fig. 6.12(a) and (b) and Fig. 6.13(a) show judgement curves for δt = 4.0 msec, 3.5 msec, and 3.0 msec, respectively.

The judged S_2 image trajectory was clearly modified by these manipulations. It is seen to have been gradually displaced towards the listener's left side and eroded by progressively earlier judgement transitions, until at δt = 3 msec it has apparently been completely biased to the left extreme position. In fact, considered in isolation, the course of the high-pitched S_2 image with reducing δt was not unlike that observed for an impulsive image when subjected to an increasing IAD favouring,in this case, the left side.

Monaural masking could be interpreted as effecting, in this instance, a reduction in the relative amplitude of the neural concomitant of the C pulse as δt was decreased. In this event, one might conclude that there would exist an effective IAD between the C pulse and the contralateral A pulse which increased with decreasing δt. The repetition period employed in these experiments was of such a magnitude that masking of the B pulse by the preceding C pulse would be expected to be of lesser importance.

It would seem initially, that the observed phenomena relating to the high-pitched impulsive image may be largely explicable in terms of monaural temporal masking. A few further experiments were conducted with the intention of clarifying this hypothesis.

Fig. 6.14(a) and (b) depict the results of interlaced experiments in which the relative levels of the B and C pulses were altered. For reference Fig. 6.13(b) shows results for the equal-level case. Fig. 6.14(a) represents the case where the level of the B pulse was reduced by 5 dB, and Fig. 6.14(b) represents the case where the level of the B pulse was increased by 5 dB. In each of the curves of Fig. 6.14 there was a decided lateral bias on the S_1 image trajectories according to the IAD applied, and the judged S_2 image trajectories may be seen to reflect the presumably reduced (6.14a) and increased (6.14b) masking of the C pulse by the preceding B pulse.

It is interesting to note the very substantial lateral bias which resulted from only a 5 dB increase or reduction in the level of the signal on one side. It has been stated earlier (Chapter 4) that in such an experiment the amplitude disparity between the B and C pulses would, if anything, be enhanced by monaural masking. Although it is not suggested that such results are directly comparable to those just described, it is perhaps relevant to consider next the curves of Fig. 6.15(a) and (b). Here, the high-and low-pitched impulsive image trajectories are shown for T = 6 msec and T = 50 msec and for IAD conditions favouring each side, in turn, by 10 dB. The listener had little difficulty in lateralizing the low-pitched impulsive image in each case, but had the impression that the high-pitched impulsive image was somewhat vague or dispersed when under the influence of the IAD.

It will be seen that, when the IAD was applied, the judgement curves were laterally biased and judgement transitions occurred nearer to ITD = 0, thus reducing the extent of the judged S_1 image trajectory. This type of behaviour is not new, but it is pertinent to compare the degree of the shift of the judgement curves here to those observed in Fig. 6.14 for the S_1 image to which the signal IAD directly applied. In the experiments of Fig. 6.14, the IAD was only 5 dB, yet the image trajectory, to which it would be expected to apply (S_1), was laterally displaced by an amount which, on the evidence of Fig. 6.15, implied a very much greater IAD. If the events are comparable, and there is no direct proof that they are, the effective enhancement of IAD in this instance is interpreted as being considerable. This observation makes even more striking the apparent absence of such effects in the case of the low-pitched impulsive image.

It is interesting to note, in passing, that the repetition period seems to be of secondary importance in the experiments of Fig. 6.15.

6.4.4 Effects of Signal Filtering: Other Impulsive Image Phenomena

Signal filtering in the simple low-pass and high-pass modes has been useful in clarifying a number of issues which have been raised in connection with the low- and high-pitched impulsive images. Certain results have suggested the possible existence of more than just two impulsive images, although the listener's impression may reflect the domination of only two images.

As precursors to a more elaborate study, a few experiments were conducted which used binaural signals band-limited to spectral regions different from those previously employed. The result of Fig. 6.16(a) was obtained with a now familiar signal configuration ($T = 6.0$ msec, $\delta t = 1.8$ msec) where the signals in both channels were band-limited to the range of 600-1500 cps. The listener reported hearing only one important image, which, when lateralized, showed judged trajectories of a new form. Here, the S_1 and S_2 image trajectories are clearly shown, but that of the S_2 image is substantially displaced forward in time. Small apparently random,variations in ITD for trajectory center crossings even up to $\pm$ 200 μsec, had been seen in earlier results obtained with this listener; however, this displacement - about 500 μsec - was so large as to suggest that this result reflected a new phenomenon.

To test the finding, another experiment was arranged using the same basic signal but with a repetition period of 40 msec. On this occasion the listener reported two impulsive images, and again described them as having low- and high-pitched characters, although the high-pitched image was different from that perceived in previous experiments. Figs. 6.16(b) and (c) illustrate the results for both images. The judged trajectory for the low-pitched image in this experiment closely resembled that of Fig. 6.16(a), hence, confirming the result; the judged trajectory of the high-pitched image, however, largely followed that of the S_1 image and was, therefore, assumed to be closely related to the high-pitched impulsive image discussed above. The results for the low-pitched impulsive image in both Figs. 6.16(a) and 6.16(b) were of particular interest, therefore, because they either represented a new phenomenon (a new image), or an interesting modification of a common form of behaviour (the S_2 image).

The immediate problem was to account for the apparent displacement of the S_2 image trajectory along the ITD axis. It was decided to approach the problem from a broad basis, and at the same time attempt to provide some information bearing on the question of the relation between monaural masking and the general spectral properties of acoustic transient signals.

A sequence of experiments was arranged which continued the pattern laid down by the experiments of Figs. 6.11, 6,12 and 6.13(a), in which the signals in both channels were identically band-limited and results were acquired at various values of δt with T held constant at 100 msec. A set of such experiments was conducted for each of the signal bandwidths, 20-400 cps, 400-800 cps, 1500-3000 cps, 3000-6000 cps, 6000-15000 cps, and for certain combinations of these bandwidths. Each bandwidth setting resulted in a modified acoustic transient waveform and Fig. 6.17 illustrates the relevant acoustic transient as measured on the artificial ear.

In this series thirty-six experiments were conducted on one listener. Not all of these results are shown as many merely confirm trends which are well illustrated by a few representative curves.

The curves of Figs. 6.18 and 6.19 depict results of experiments using binaural signals bandlimited 20-400 cps (effectively 400 cps low-pass). Figs. 6.18(a) and 6.18(b) show that at δt values of 12 msec and 6 msec both S_1 and S_2 image trajectories may be clearly observed. Figs. 6.19(a), (b) and (c) illustrate that the major effect of further reducing δt was to alter the relative importance of the S_2 image as indicated by the ranked judgements; however, at $\delta t = 1.0$ msec the S_2 image was apparently not reported as an independent entity, but the modification of the judgement trajectory in the ITD interval between pulses B and C suggests that it was still exercising some influence on judgements. These results are similar to those of Figs. 6.10(b) and (c), where signals were bandlimited 600 cps low-pass.

Similarly, Figs. 6.20 and 6.21 illustrate judgement curves for $\delta t = 12$ msec, 6 msec, 3 msec, 2 msec and for binaural single-pulse trains, when the signals were bandlimited 400-800 cps. Once again the S_1 and S_2 image trajectories are clearly seen for $\delta t = 12$ msec and 6 msec; however, at $\delta t = 3$ msec and 2 msec the S_2 image is much less in evidence. Curious features of these curves are the conspicuous appearances of subsidiary image trajectories spaced about 1 to 1.5 msec from the major S_1 and S_2 image trajectories. These subsidiary image trajectories are apparent also in Fig. 6.21(c) where the C pulses have been omitted from the signal; it will be remembered that a similar phenomenon has been observed before and was discussed briefly in Chapter 4 (see Fig. 4.29).

Continuing the series, Figs. 6.22 and 6.23 illustrate results of experiments using signals bandlimited 800-1500 cps; Figs. 22(a) and (b) for δt = 12 msec and 3 msec, and Figs. 6.23(a) and (b) for δt = 6 msec and for binaural single repetitive transients. As before, the S_2 image was seen to become progressively less important in the judgements as δt was reduced below about 6 msec (Fig. 6.23(a) , and at δt = 3 msec (Fig. 6.22(b) was not at all in evidence. Fig. 6.23(b) illustrates results using A and B pulses only. The subsidiary image trajectories were once more conspicuous, here positioned somewhat closer to the S_1 and S_2 image trajectories. Spacings in these experiments were seen to be about 0.6 - 1.0 msec.

Figs. 6.24 and 6.25 depict results for δt = 12 msec, 6 msec and δt = 3 msec when the signals were bandlimited 1500-3000 cps. At δt = 12 msec and 6 msec both S_1 and S_2 image trajectories were clearly seen, but at δt = 3 msec the S_2 trajectory was apparently absent. The existence of subsidiary images again may be inferred from the curves. Here the separation between main and subsidiary image trajectories were in the range 0.5 - 0.8 msec, with one example (Fig. 6.25a) where the spacing reached 1.2 msec.

Results using signals bandlimited 3000-6000 cps were very similar to those which used signals bandlimited 6000-15000 cps. The latter findings have been shown earlier in Figs. 6.11, 6.12 and 6.13(a).

From these results a number of important points have emerged, viz.,

(1) binaural transient signals of different spectra (i.e., waveform) may give rise to impulsive images which have broadly similar lateralization characteristics under the influence of ITD. Differences in detail, however, may be observed and certain of these differences appear to be systematically related to signal bandwidths.

(2) both S_1 and S_2 image trajectories were seen for all signals used when δt = about 6 msec. For signal bandwidths including the extreme low-end of the audible spectrum, however, the two trajectories were observed at δt values as low as 1.2 msec (Fig. 6.10b). With the exception of these very low-pitched images, the effects of monaural masking appear to be broadly similar in each case (Figs. 6.18 - 6.25 and 6.11 - 6.13). In each of the sets of experiments reported above, the importance of the S_2 image was apparently reduced as δt was reduced. This could be inferred either from the progressive elimination of the judged S_2 trajectory in most of these experiments or from the altered judgement ranking in others (e.g., Figs. 6.18, 6.19 and 6.10). In either case δt = 6 msec appeared to be the point at which this effect became important. Observation of the results for progressively higher pitched experimental signal bandwidths revealed an interesting trend wherein the judgement transitions of the lower-pitched S_2 images appeared to be biased towards the right side as δt was reduced (e.g., Fig. 6.18) and judgement transitions of the higher-pitched S_2 images appeared to be biased towards the left side as δt was reduced (e.g., Fig. 6.23).

(3) subsidiary image trajectories were seen to appear regularly with signals of certain bandwidths. The spacings between the subsidiary image trajectories and the main S_1 or S_2 image trajectories were seen to bear a simple relationship to the signal spectrum and/or transient waveform. These images were most readily apparent when the signals were bandlimited to frequencies below about 3000 cps, and appeared not to be restricted to those experiments which used repetitive single-versus double-pulse trains.

(4) the apparent forward displacement in ITD of the S_2 image trajectory that was observed in experiments of Fig. 6.16 now appears to be a phenomenon not involving the S_2 image alone. Figs. 6.22 and 6.23 depict results of experiments using signals having bandwidths similar to those employed in the experiments of Fig. 6.16. It is clear that at δt values less than 6 msec the S_2 image was of little importance in the judgements, and that the judged trajectory interpreted earlier as being that of the S_2 image (where $\delta t = 1.8$ msec) was almost certainly that of a subsidiary image, perhaps influenced somewhat by the presumably laterally-biased and similar-sounding S_2 image.

Having observed judged trajectories of the images which arise from approximately octave-band filtered acoustic transients, it was considered to be of interest next to observe judged trajectories of images which arise from binaural acoustic-transients filtered so as to include multiple octave-bands.

Experiments were conducted which used binaural signals bandlimited to 400-1500 cps, thus including the frequency bands 400-800 cps and 800-1500 cps.

For the situations in which δt = 6 msec (Figs. 6.20b and 6.23a) the judged S_2 image trajectories for signals of these bandwidths were significantly different. The result of the 'combined' experiment (not shown) was virtually identical to that of Fig. 6.23 (a) indicating, perhaps, the stronger influence of spectral components in the band 800-1500 cps. It would appear to be significant that the similarity extended even to the precise positioning of the subsidiary image trajectory.

Next, the same experiments were conducted using signals bandlimited 20-1500 cps. As in the previous experiment the listener judged the dominant image, and did not attempt to especially seek out the low-pitched impulsive image. Here, again, the result (not shown) was almost identical to that of Fig. 6.23(a) with the effects of the subsidiary image somewhat reduced, but generally similar.

Another relevant and interesting sequence of results was that obtained using experiments in which the right-ear signal was unfiltered and the left-ear signal was bandpass filtered. In these experiments, as in experiments of similar type described earlier, the listener reported a 'residue' image positionally fixed at the right-extreme. This made judgements of relative importance very difficult but in spite of this the results obtained by this method were very consistent and each closely resembled the appropriate experimental result where both signals were correspondingly bandlimited. An example of this type of experiment is shown in Fig. 6.26(a), where the left-ear signal was bandlimited to 6000-15000 cps and δt = 4.5 msec. This result should be compared with that of Fig. 6.11(b); when allowance is made for the displacement along the

ITD axis, due presumably to the dissimilar acoustic transient waveforms in each channel, the results are seen to be closely similar. Fig. 6.26(b) illustrates the result of an experiment of similar type, wherein the left-ear signal was bandlimited to 800-1500 cps and $\delta t = 6.0$ msec; see Fig. 6.23(a) for comparison.

In each of these latter experiments the subsidiary images were not so conspicuous as in the earlier results.

6.4.5. Effects of Selective Masking

Experiments utilizing selectively filtered acoustic transients have been considered in some detail, and have proved useful in further elucidating a number of impulsive image phenomena; however, certain disadvantages of the technique restrict its usefulness. The modification of the acoustic transient waveform by filtering introduced a complicating factor in certain experiments where shifts of judgement trajectories along the ITD axis, possibly as a result of IAD, were important information. Secondly, and even more important, although mechanical cochlear activity is known to be broadly a function of the signal spectrum it is virtually impossible directly to know exactly how the cochlear partition moves in response to a given acoustic transient.

Even though the addition of masking noise to the signal may avoid the first objection, this technique also is subject to the second limitation; however, whereas low-pass filtering of the acoustic signal cannot completely eliminate motion of the cochlear partition in the basal region, the addition of sufficient. high-pass filtered noise to the signal would be expected to effectively mask neural activity in this region, whether the signal was filtered or not.

As a simple introduction to the method, a few casual experiments were conducted to ascertain the effects of masking noise on the low-and high-pitched impulsive images normally perceived with unfiltered transient signals. Listeners adjusted the low-or high-pass cutoff frequencies of the filters through which the masking noise was passed, and reported the settings at which changes in character of the low-or high-pitched images were just perceptible. In general, these cutoff frequencies fell in the range 1000-2500 cps, closely corresponding to those reported in similar experiments where the signals were filtered (see Section 6.4.1). As expected the high-pass noise eliminated the high-pitched image, and low-pass noise eliminated the low-pitched image.

More instructive were experiments of a familiar type, $T = 6$ msec, $\delta t = 1.8$ msec, using unfiltered , equal-level transients with low-or high-pass noise added to the signals in either or both ears. Figs. 6.27 and 6.28 depict results of three interlaced experiments in which the low-and high-pitched impulsive images were simultaneously lateralized (no masking noise) and then were lateralized separately (high-and low-pass masking noise, respectively). Fig. 6.27 illustrates the now familiar results for the low-and high-pitched impulsive images, as a basis for direct comparison.

Fig. 6.28(a) illustrates the result when high-pass filtered noise - cutoff frequency = 1500 cps - was applied with the signal in the right ear. Here, the listener reported the high-pitched image as being fixed at the left-extreme position; the low-pitched image, however, was clearly perceived and was lateralized in a fashion which closely resembled the curves of Fig. 6.7 where the low-pitched image was independently tracked after a sequence of preparatory experiments.

Fig. 6.28 (b) illustrates the judged trajectory of the high-pitched impulsive image perceived when low - pass masking noise, was added to the right-ear signal. Even though it cannot be assumed that no masking takes place in the high-frequency (basal) portions of the cochlea, the curve is very similar to that of Fig. 6.27 (b), and other comparable results for high-pitched impulsive images.

Results of these experiments were closely similar when masking noise was added to the left-ear signal alone, or when, it was added to both left- and right-ear signals.

Continuing from the experiments of Figs. 6.27 and 6,28, two further experiments were conducted. In these the cutoff frequency of the high-pass masking noise in the right-ear channel was raised to 3000 cps and then to 6000 cps. With the masking noise filtered to 3000 cps high-pass, the listener reported only a single roving image, which,when lateralized,gave results similar to those of Fig. 6.28(a); however, when the cutoff frequency was raised to 6000 cps the listener reported a second image of high-pitched character. Judgement curves for both images are illustrated in Fig. 6.29, and they are seen to be virtually identical to the curves of Fig. 6.27.

This is a significant result, for it suggests that when both high- and low-pitched impulsive images are perceived, judgements of one may be influenced by the presence of the other, unless special procedures are undertaken to particularly direct the listener's attention to each image individually. This would seem to apply mainly to the influence of the high-pitched image on positional judgements of the low-pitched image. In experiments where the high-pitched image was lateralized it was seen to have been judged similarly in each case; however, the judged trajectories for the low-pitched impulsive image were seen to adopt one of two patterns - either showing very clearly, and almost completely, the S_1 and S_2 image paths (Fig. 6.28a), or demonstrating the reduced effects of the S_2 image (Figs. 6.27 and 6.29), as were illustrated and discussed in Chapter 4. The former pattern was previously seen on occasions where procedures were employed to direct the listener's attention to the low-pitched image, and where the listener was sufficiently experienced to recognize a split-image situation, for the S_2 image was accompanied by an important S_1 image positioned at the right-extreme position (Fig. 6.7). In these earlier experiments extreme low-pass filtering of the acoustic transients was also often sufficient to enable unrestrained lateralization of the S_2 image at small δt values, but cutoff frequencies as low as 400-600 cps were necessary before the effect could be observed, (Figs. 6.10 and 6.19). Here, however, masking of spectral components above 3000 cps was sufficient to enable the low-pitched S_2 image to be clearly lateralized; although when sufficient of the higher spectral components were unmasked, the high-pitched impulsive image was independently perceived and immediately the judgement pattern of the low-pitched image was observed to alter (Fig. 6.29).

It was apparent from these results that unilateral masking of the acoustic signals did not appreciably alter the ITD values at which the judged trajectories crossed the centre axis. To clarify this important matter a few experiments were conducted in which masking noise of different levels and cutoff frequencies was employed.

The first of these experiments used binaural single-pulse trains to which had been added, in the right-ear channel, masking noise high-pass filtered with cutoff frequencies of 1500 and 600 cps. Fig. 6.30(a), (b) and (c) depict results of these interlaced experiments accompanied by results of a control experiment using unmasked signals conducted at the same time. For reasons of brevity data was not taken over the full signal period. As in the other masking experiments the level of the noise was established such that wideband noise completely masked the wideband transient signal. As was expected the listener reported a 'residue' image of high-pitched character at the left lateral extreme position. The listener reported little difficulty in independently lateralizing the low-pitched image when the masking noise was 1500 cps high-pass filtered, but expressed some difficulty when the noise was 600 cps high-pass filtered. In the latter situation, the low-pitched image was reported as being very faint and dispersed, often not heard at all. The curves show closely similar ITD values for centre crossings, even in Fig. 6.30(a) where judgements were only occasionally reported. An interesting aspect of Fig. 6.30(b) is the indication of a subsidiary image, spaced about 1msec from the S_1 trajectory.

A comparison of these curves with those of Figs. 6.3 and 6.4(a) is instructive. The latter curves relate to experiments which used, in one channel, filtered transients having spectra similar to the unmasked bandwidths of the transients in these experiments.

A result typical of experiments which used transients of different bandwidths in each channel, is shown in Fig. 6.31(a). Here, an extreme example was chosen where the left-ear signal was 400 cps low-pass filtered, and the right-ear signal was unfiltered. The trajectory is seen to cross the centre axis about 1.1 msec from ITD = 0; this would be anticipated from consideration of the acoustic waveforms used - see Fig. 6.17. When high-pass filtered masking noise was added to the right-ear signal, no change could be observed in the judged trajectory of this low-pitched image. Figs. 6.31(b), (c) and (d) illustrate curves for experiments using high-pass noise with cutoff frequencies of 1500, 800 and 20 cps. It will be noted that even when the masking noise was wideband and of sufficient intensity to monaurally render the right-ear signal inaudible the image was still lateralized. This was attributed to the fact that in order to produce a clear low-pitched image the signal level in the filtered, but unmasked, left channel had to be increased by 6 dB. The listener, not surprisingly, had considerable difficulty in finding and lateralizing the image when the masking noise was wide-band, and reported the impression that the image tended to move about but with a discernable central tendency. Application of the masking noise to the opposite channel made no change in these results.

A further sequence of experiments was carried out which was intended to demonstrate the effects of masking noise applied at various levels. These experiments used repetitive single-pulse trains; equal-level, T = 6 msec, unfiltered. Unfiltered masking noise was added to the right-ear signal at levels ranging from that at which the noise was barely audible in the presence of the signal, to about 5 dB below that for complete masking of the signal - a span of about 20 dB, at the signal levels employed. In these experiments the low-and high-pitched impulsive image, as well as the second harmonic tonal image, were lateralized.

At the lowest level of masking noise, the judged trajectories (not shown) of all three images were identical to those shown earlier for the comparable experiment with no masking noise, shown in Fig. 5.8 and 5.9 (the broken curves). Results were taken with the noise level increased by 10 dB and then by 20 dB, at which level the right-ear signal was almost completely masked. The increase of 10 dB was apparently sufficient to mask the high-pitched image contributions from the right ear because this image was perceived thereafter to be fixed at the left extreme position. The low-pitched impulsive image and the second harmonic tonal image trajectories were unaltered; nor were they seriously modified by the 20 dB increase in masking noise level. Careful comparison of curves obtained for the lowest and highest levels of masking noise indicated a possible displacement of the curves towards the left side, but the amount of the displacement was very small indeed, corresponding to a shift of centre-crossing ITD of 50 µsec, which in these experiments is not of great significance.

6.5.0 Discussion

In general, it appears that the technique based upon the independent identification and lateralization of the various impulsive image components are more powerful than previous methods (e.g., centering) in elucidating the complex perceptions which arise from binaural presentations of acoustic transient stimuli. The accumulated background of experimental information relating acoustic image lateralization judgements to ITD and IAD parameters for a variety of relatively simple signal configurations has proved to be of valuable assistance in interpreting results of the more complex experiments. Perhaps the most universally useful signal has been the repetitive pulse - versus pulse-pair configuration.

It now seems possible to consider some general aspects of neurophysiology, make some preliminary inferences as to the sources of neural activity giving rise to the impulsive images, and to speculate as to how this activity interacts binaurally. In the course of the following discussion attention will be drawn to some of these important points; however, the major discussion on these issues will be left to the following chapter which deals specifically with some relevant neurophysiological matters.

The results of the experiments in this series may be summarized as follows:

1. Studies of impulsive image phenomena have made it possible to come to some general conclusions as to the mechanisms which underlie the perceptions involved. There would appear to be substantial evidence to support the view that the low-pitched impulsive images are associated with spectral components of the transient signal which lie below about 1500 cps. The high-pitched impulsive images appear to be associated largely with spectral components above about 1500 cps. In fact, it has been possible to be more specific and to relate these images to neural activities apparently arising from specific portions of the cochlea; the low-pitched impulsive image appears to be associated

with neural activity in the apical region, and the high-pitched impulsive image appears to be associated with neural activity in the basal region. The characters of the relevant images, as exemplified by the lateralization properties, were altered when synchronous motion of the basilar membrane in either of these regions was effectively reduced. The effects as reflected in the judged trajectories were similar when the signals were selectively bandpass filtered and when the signals were masked with appropriately bandpass filtered random noise. Where the modification applied to one signal only, the impulsive image associated with frequencies within the stop-band of the filtered signal or within the pass-band of the filtered masking noise, was perceived to be firmly displaced to the side of the unaltered signal. If the cutoff frequency of the filter, applied either to the signal or to the masking noise, was in the region of 1500 cps, the commonly-identified low-or high-pitched impulsive image was manipulated. If, however, the cutoff frequency was set to any other value, the pitch-like character of these images may have been perceived to alter, and in certain instances, other impulsive images, either fixed or roving, may have been identified.

It seemed to be generally true that where the effective bandwidths (i.e., the portion of the signal bandwidth which is neither attenuated by filtering nor masked by noise) of the binaural transient signal overlapped,an impulsive image was perceived which could be judged to move with ITD. This roving image was reported as having a pitch-like character which could be related to frequencies within the common effective bandwidth. Spectral components outside the effective bandwidth which was common to both signals apparently did not enter into binaural fusion, of if fused were presumably subject to considerable IAD. In any event, the impulsive image of relevant pitch was observed to remain at the lateral extreme position.

Where bandwidths of the signals in both channels did not overlap (or were overlapping only to the extent that the attenuation slopes of filters in each channel were such that some slightly supra-threshold components were common to both channels) the image phenomena were apparently very inconsistent. Under these conditions listeners' reports varied considerably; some heard no roving image, others reported only a "sensation," others could not be certain. None reported hearing images of the type normally perceived that clearly altered position with ITD.

2. It is argued that monaural masking effects have been substantially clarified. These effects are apparently not equally operative in the mechanisms which underlie the low-and high-pitched impulsive image phenomena; low-pitched impulsive images were apparently little affected, whereas high-pitched impulsive images were clearly influenced in a systematic fashion.

3. Views, outlined in earlier chapters, on the basis of judgements in split-image situations have been confirmed. An amendment is proposed, however, which includes situations in which images of closely similar (i.e., impulsive) character are simultaneously perceived in different positions. In experiments using repetitive single - versus double-pulse trains, the perception of the high-pitched impulsive image was seen clearly to influence positional judgements of the low-pitched impulsive image. See Figs. 6.27, 6.28 and 6.29.

It was possible, however, to apparently minimize such interaction by special, but simple, training procedures which directed the listener's attention to the low-pitched image (Fig. 6.7).

4. The general waveform-dependence of the impulsive images which was previously hypothesized has been confirmed. There is evidence to suggest, however, that simple consideration of acoustic transient waveforms may not be adequate. Results of experiments using dissimilarly bandpass filtered acoustic transients in each channel were particularly instructive in this respect. Filtering of the signals altered the acoustic transient waveforms. Where the filtered signals had a substantial bandwidth in common, a well-defined impulsive image was generally perceived which, when lateralized, was observed to cross the centre axis at ITD different from zero. This ITD could be closely predicted from observation of the relevant acoustic transient waveforms and appeared to closely equal the ITD required to bring the initial peaks of the measured transients into alignment. The unequal filtering of the binaural acoustic transients appeared not to introduce any important amount of IAD. This could be deduced from a number of results, but perhaps most convincingly from the results of Figs. 6.5(b) and 6.6(a) and (b) where the general forms of the judgement curves are unchanged, but are merely shifted along the ITD axis by the signal manipulations. This signal configuration is believed to be a particularly sensitive indicator of any IAD parameter changes.

It is interesting to consider these results in the light of current hypotheses of basilar membrane motion and associated neural activity. An hypothesis which has found wide acceptance is that proposed by

Flanagan (1962) wherein neural firings are associated with unipolar peaks of basilar membrane displacement. It has been stated from time to time in this chapter that, in experiments using different acoustic transients in each channel, there was a close agreement between the ITD for a judged centered image and the ITD between initial peaks of the measured acoustic transients. It could be suggested that the images resulted from the binaural interaction of neural firings directly associated with these unipolar peaks; however, the explanation is clearly incomplete at that.

It is believed that basilar membrane responses to acoustic signals behave as if the signals undergo a form of progressive low-pass filtering (Flanagan, 1960, 1962). On this hypothesis, unfiltered transients would be expected to be progressively modified as they proceed towards the apex of the cochlea, while low-pass filtered transients would be expected to be transmitted, substantially unmodified, down to that region of the basilar membrane maximally responsive to the highest frequencies contained in the transient, and thereafter be modified. If neural information from various parts of cochlea is important in the creation of binaural images, then neural transductions of basilar membrane motion would be expected to reflect, in some way, the characteristics of this low-pass filtering process. There is strong evidence to suggest that this highly simplified view may be a reasonable first approximation: Davis (1957), Flanagan (1962), Teas, et al. (1962). The issue will be considered in more detail in the following chapter, but assume for the moment that the above statements are a fair representation of the facts.

There were two experimental observations which cast some doubt on any hypothesis which considers only the only the acoustic transients waveform as measured at, say, the oval window. These experimental findings were: first, the appearance of subsidiary image trajectories in earlier experiments (Chapter 4) where the signals were wideband and unmasked and, second, the appearance of similar trajectories when frequencies above 1500 cps were masked with filtered random noise, Fig. 6.30(b). Both of these observations were inexplicable in terms of measured acoustic transient waveforms, but it will be seen that they could be accommodated by consideration of basilar membrane motion in the region maximally responsive to about 1000 cps. According to Flanagan (1962) and Harris, et al. (1963) wide-band acoustic transients would be expected to cause displacement maxima in the 1000-1500 cps region of the basilar membrane. This region might be expected to show repetitive unipolar displacements at temporal intervals which could be directly associated with the separation in ITD (about 1 msec) of the main and subsidiary image trajectories in the above mentioned experiments.

The psychophysical findings of Flanagan (1962) are also relevant to the present issue. Flanagan reported experiments which used binaural unfiltered transients which were binaurally masked with the uncorrelated high-pass filtered noise. Listeners adjusted the ITD to bring images to the central position. For cophasic transients, image centerings were found at ITD = 0 and at other ITD values which agreed reasonably with the relation $\pm 1/2f_c$, where f_c = cutoff frequency of the high-pass filters. Flanagan argued that the dynamic properties of the middle ear and the basilar membrane were such that for cutoff frequencies below 1500 cps the maximum coherent neural activity would arise from that portion of the membrane which was just beyond the distal limit of the masking noise. His findings would appear to support this view.

It is interesting to apply this same principle to an examination of experiments using bandpass filtered transients.

It has been shown that binaural acoustic transients filtered similarly at both ears gave rise to impulsive images when signal bandwidths were widely varied. Judged trajectories of these images often showed the effects of monaural masking and also often implied the existence of subsidiary images. Where subsidiary images were apparent, the spacing between the main image trajectories and the subsidiary image trajectories appeared to bear a simple relation to the higher cutoff frequency of the bandpass filter. This relation was observed to be largely unmodified by alterations of the lower cutoff frequency. Superficially, the resemblance between this result and the results of Flanagan (1962) is striking. The implication would appear to be that in this respect low-pass (say) filtering of the acoustic signal has largely the same effects as masking of a broadband signal with high-pass masking noise. The results of experiments using symmetrically bandpass-filtered transients were clear enough, but without further study it was not obvious as to which region of the cochlea contributed to the perception of either the main or the subsidiary images. The situation was hardly resolved by the observation that the unipolar peaks of the measured acoustic transient waveform were repetitive at approximately the period of the highest spectral component (i.e., the higher cutoff frequency). On the hypothesis presented above it would be expected that basilar membrane displacements, from the base to some point near that maximally responsive to the higher cutoff frequency, would be dominated by the temporal features of the acoustic transient waveform. Therefore, in contrast to the previous cases, which used wideband transients, there would be no unique narrow region of the membrane with the required displacement characteristic. In fact, the results could be described on the basis of neural timing information arising from the basal region alone.

It would seem likely that bandpass filtered signals give rise to maximum displacements of the basilar membrane in those regions which are normally selective to frequencies within the pass-band; however, it is an observed fact that all such signals cause some activity in the basal turn of the cochlea. Since the basal turn includes about half the neural population of the cochlea (see for example Deatherage, et al., 1959) it is possible that even relatively slight motion of the basilar membrane may cause appreciable volleys of neural activity from this region. A number of experimental results suggest, however, that such effects may not be of direct importance in the creation of the impulsive images in question. For example, the results of Fig. 6.31 and other similar results demonstrated quite convincingly that high-pass filtered masking noise did not seriously interfere with the perception of the impulsive image which arose from unilaterally low-pass filtered signals, thus implying that the image was the result of neural activity in more apical regions.

Further evidence bearing on this matter comes from Fig. 6.26(b) where subsidiary image trajectories, although less pronounced, may be seen near the S_1 image trajectory. In spite of the drastically modified acoustic waveform in the right channel the subsidiary image trajectories in this result could be interpreted as being similar to those of Fig. 6.23(a), 6.23(b) and others, which were obtained in experiments using identically filtered transients in each channel.

Less direct evidence comes from experiments dealing with monaural masking. It has been shown that images attributed to neural activity in the basal turn have been more susceptible to monaural masking effects than low-pitched impulsive images believed to be the result of neural activity in the apical region. It appears from these findings that monaural masking effects in the basal region are such as to prevent completely the lateralization of a subsidiary image spaced as closely to the main image trajectory as 1 msec.

These pieces of evidence seem to confirm the impression that neural activity arising from the basal turn of the cochlea was not of fundamental importance to the formation of impulsive images arising in experiments using signals bandpass filtered so as to include only those frequencies below about 1500 cps. The preliminary conclusion which is drawn from such experiments is that neural activity in the apical regions of the cochlea was largely responsible for the dominant impulsive images which were lateralized and that the appearances of subsidiary images were, in an important sense, related to the dynamic responses of the basilar membrane to the filtered acoustic signals.

Implicit in these arguments are certain factors relating to the mechanisms of binaural interaction. It would seem that not only may the various images be associated with neural activity arising in various portions of the cochlea, but that these images may arise from cross-comparisons of neural activity from, in particular, corresponding regions of the contralateral cochleae. The most direct evidence to this effect comes from experiments in which high-pass filtered masking noise was applied to either or both channels with (in this respect) no important modification of judged trajectory of the low-pitched impulsive image from that determined with no masking noise.

A few further and rather more specialized experiments were necessary to clarify certain of these issues. The following chapter contains descriptions of some of these experiments and a continuation of the arguments initiated above. Discussion of related research by other workers will be reserved until that time.

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C = 1:1:1   Filtered 20-1500 cps, B Pulses only.

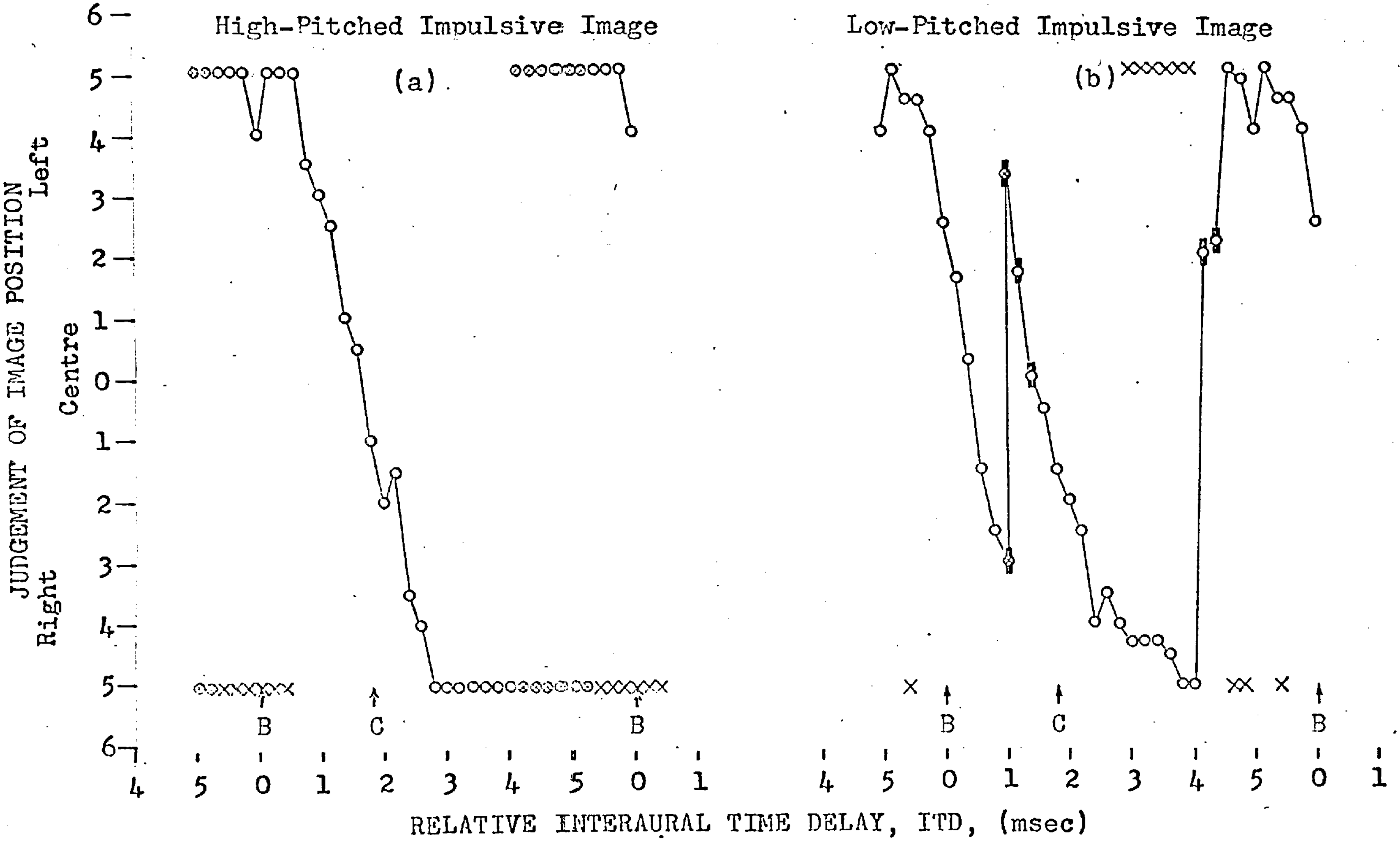


FIGURE 6.1

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec   δt = 1.8 msec
A:B:C = 1:1:1   Filtered 20-1500 cps, C Pulses only.

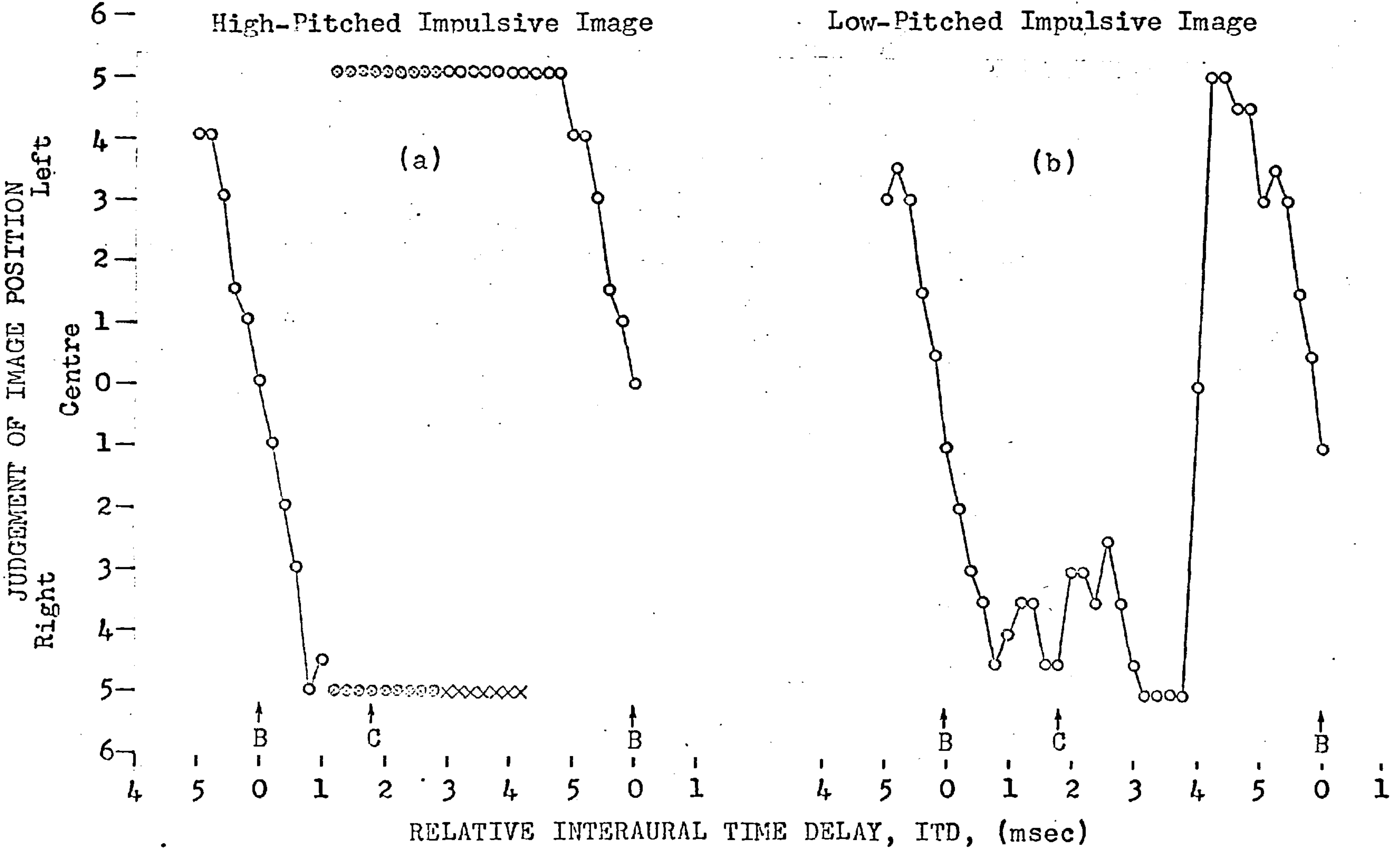


FIGURE 6.2

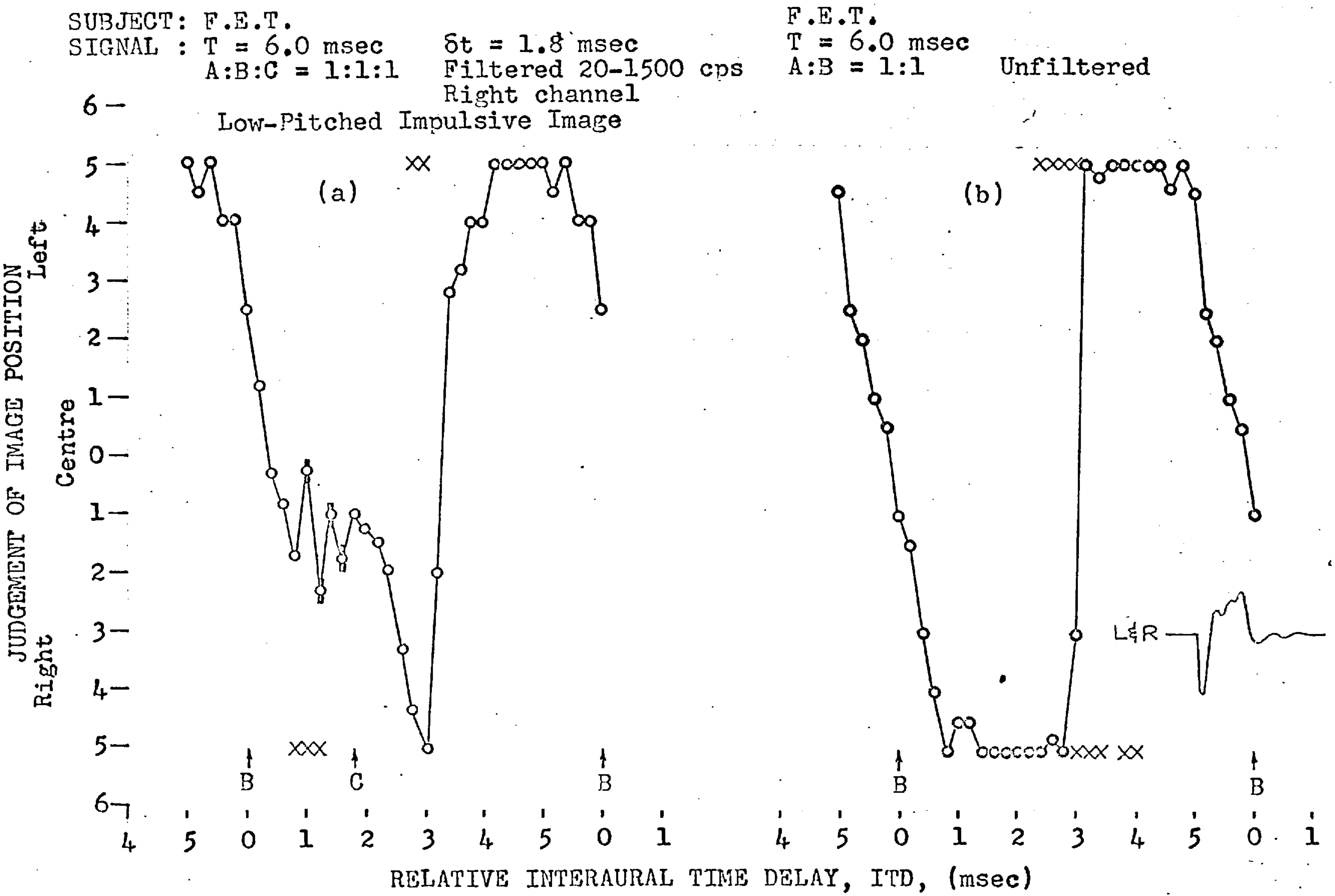

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec
A:B:C = 1:1:1
δt = 1.8 msec
Filtered 20-1500 cps
Right channel
Low-Pitched Impulsive Image
F.E.T.
T = 6.0 msec
A:B = 1:1
Unfiltered
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
B
L&R
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.3

②

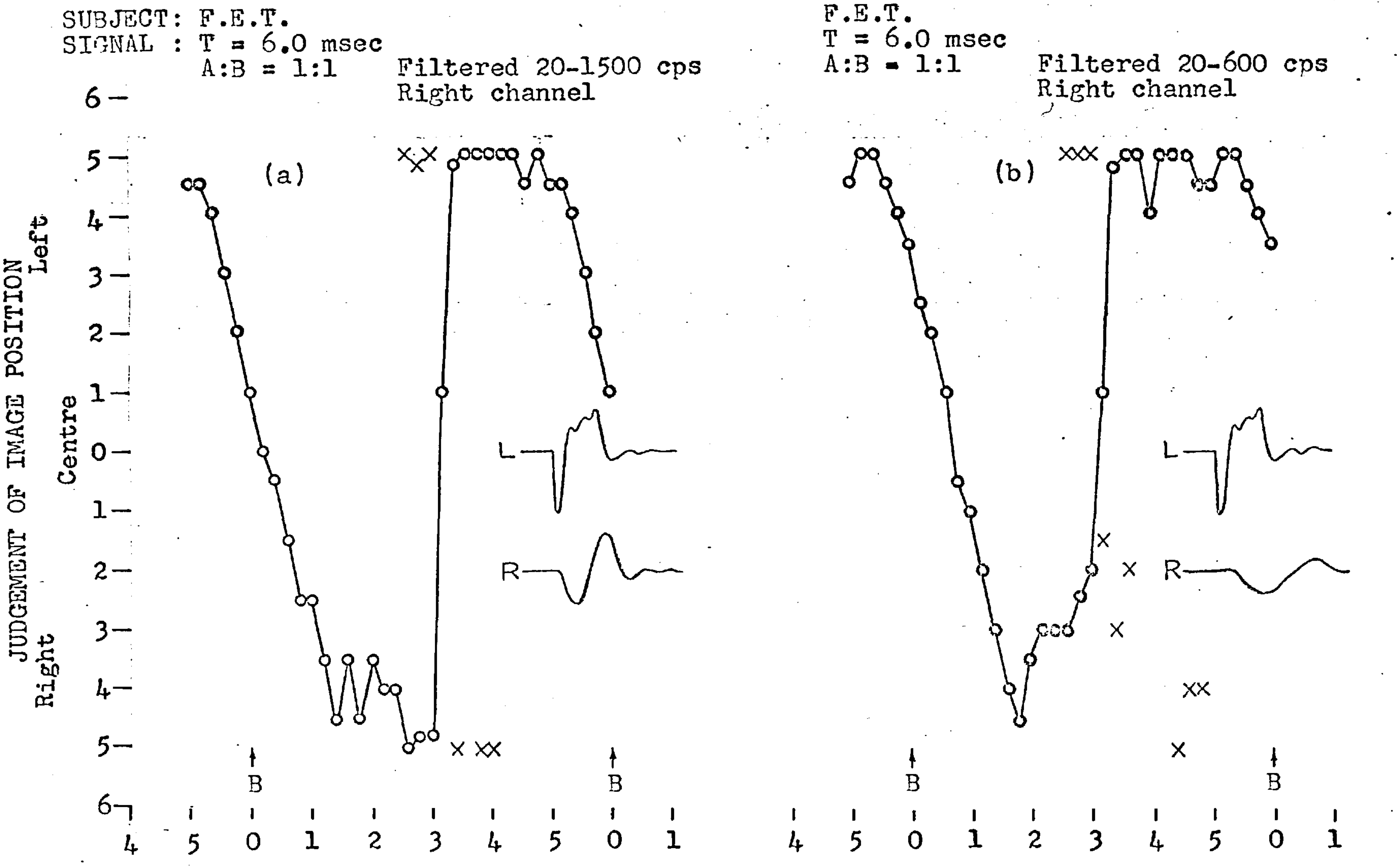

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec
A:B = 1:1
Filtered 20-1500 cps
Right channel
F.E.T.
T = 6.0 msec
A:B = 1:1
Filtered 20-600 cps
Right channel
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
L
R
B
B
B
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.4

②

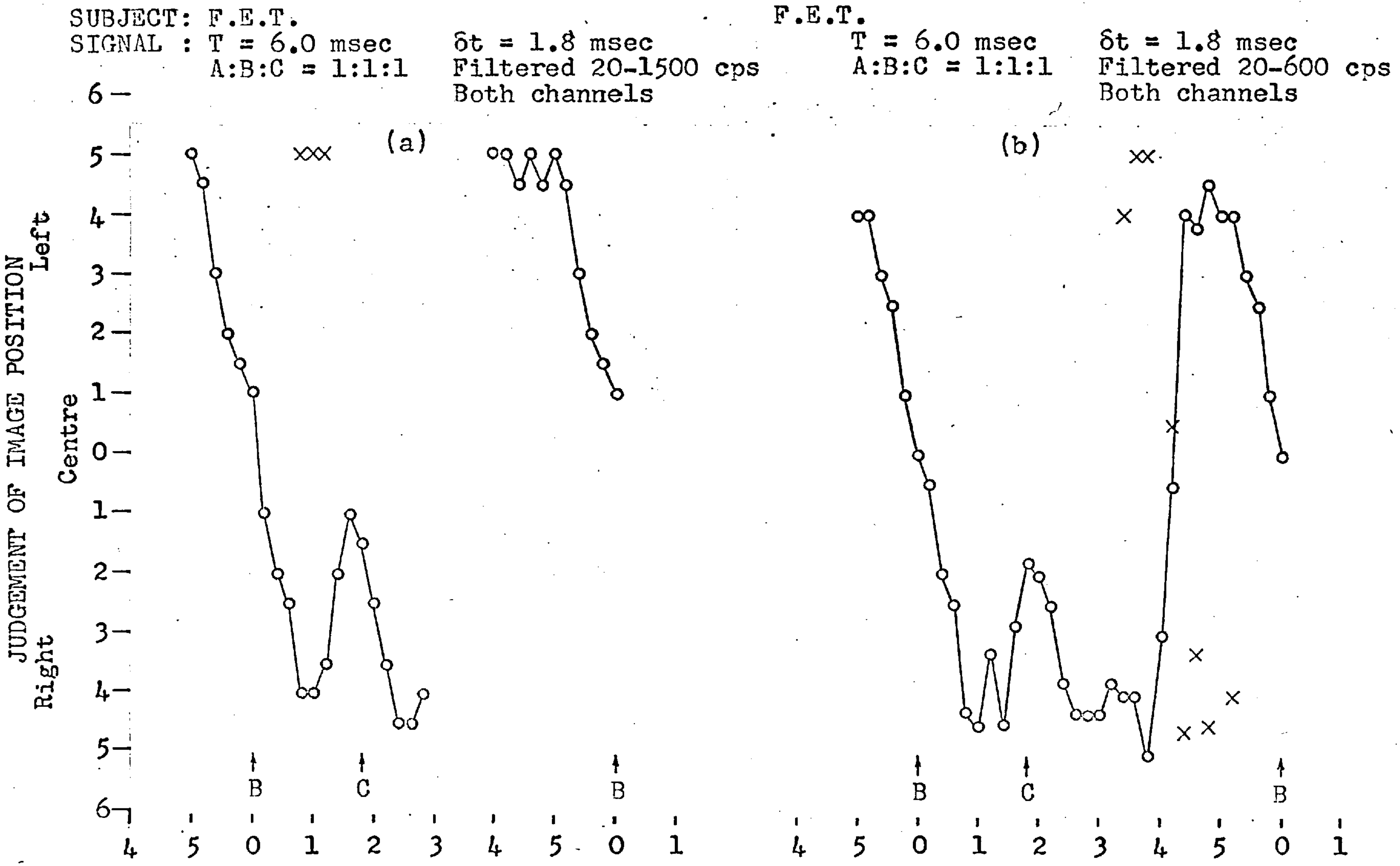

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec
A:B:C = 1:1:1
δt = 1.8 msec
Filtered 20-1500 cps
Both channels
F.E.T.
T = 6.0 msec
A:B:C = 1:1:1
δt = 1.8 msec
Filtered 20-600 cps
Both channels
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.5

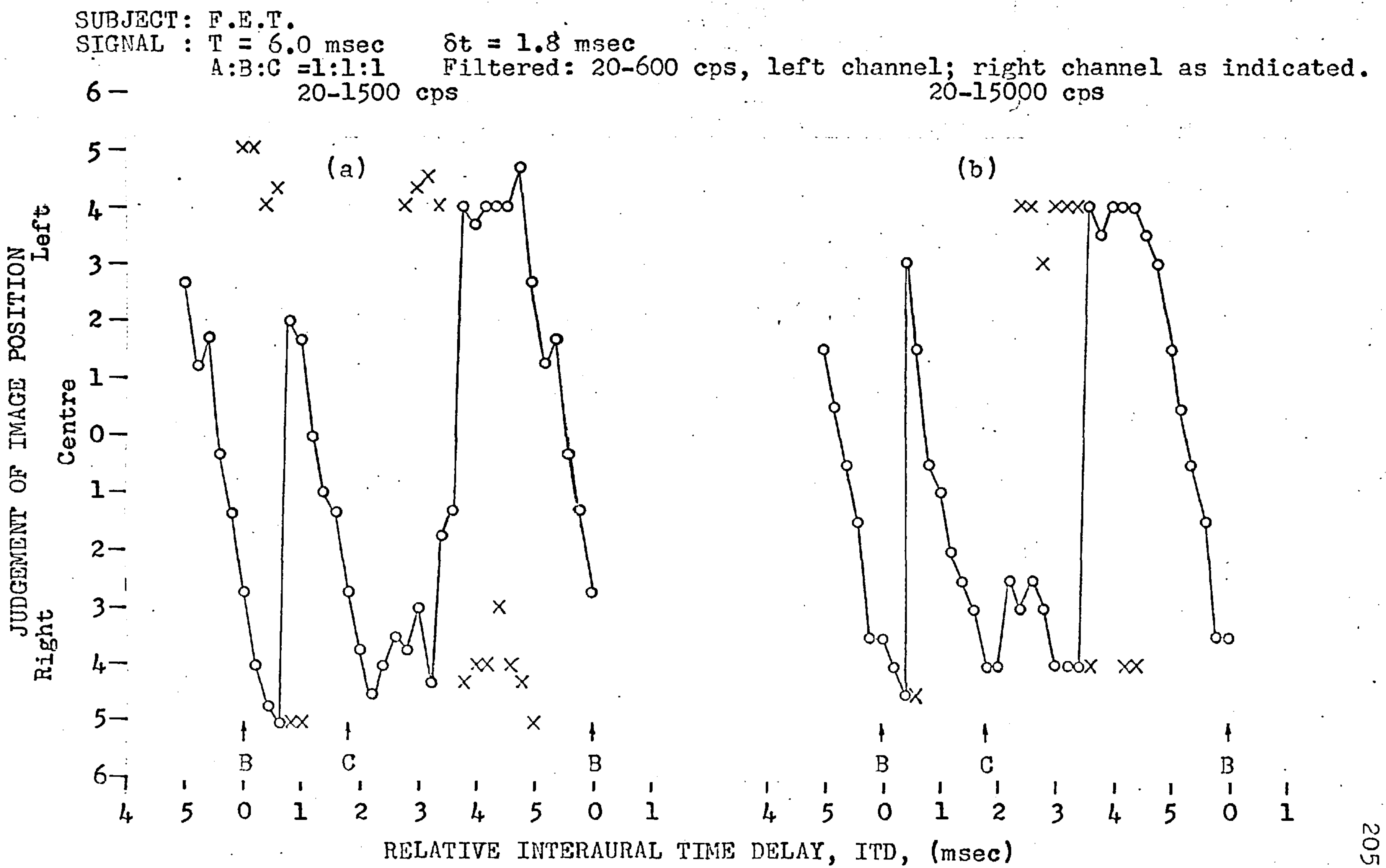

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec δt = 1.8 msec
A:B:C =1:1:1 Filtered: 20-600 cps, left channel; right channel as indicated.
20-1500 cps
20-15000 cps
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.6

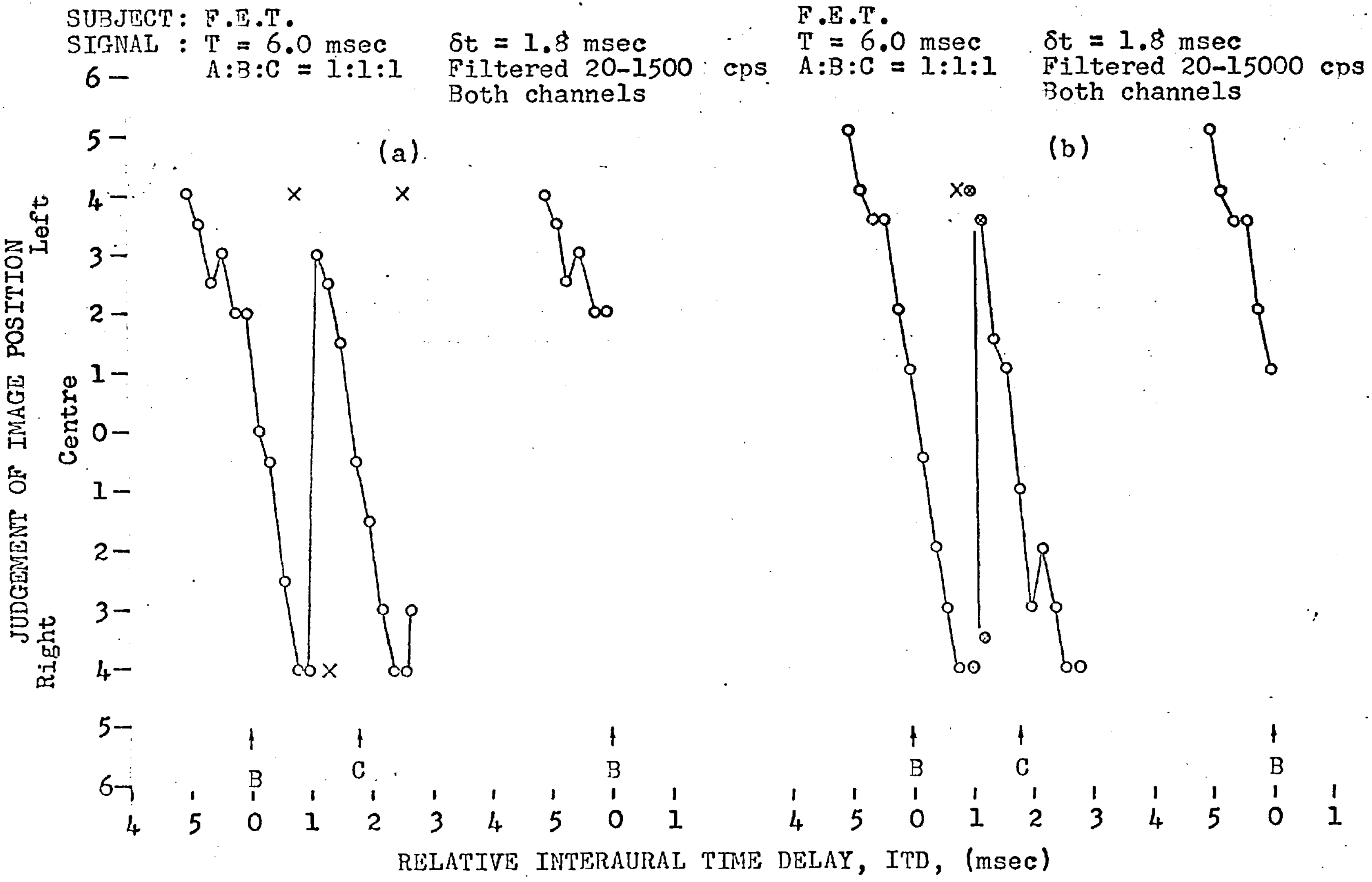

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec
A:B:C = 1:1:1
δt = 1.8 msec
Filtered 20-1500 cps
Both channels
(a)
F.E.T.
T = 6.0 msec
A:B:C = 1:1:1
δt = 1.8 msec
Filtered 20-15000 cps
Both channels
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.7

SUBJECT: F.E.T.
SIGNAL : T = as indicated δt = 2.4 msec
A:B:C = 1:1:1 Unfiltered

High-Pitched Impulsive Image

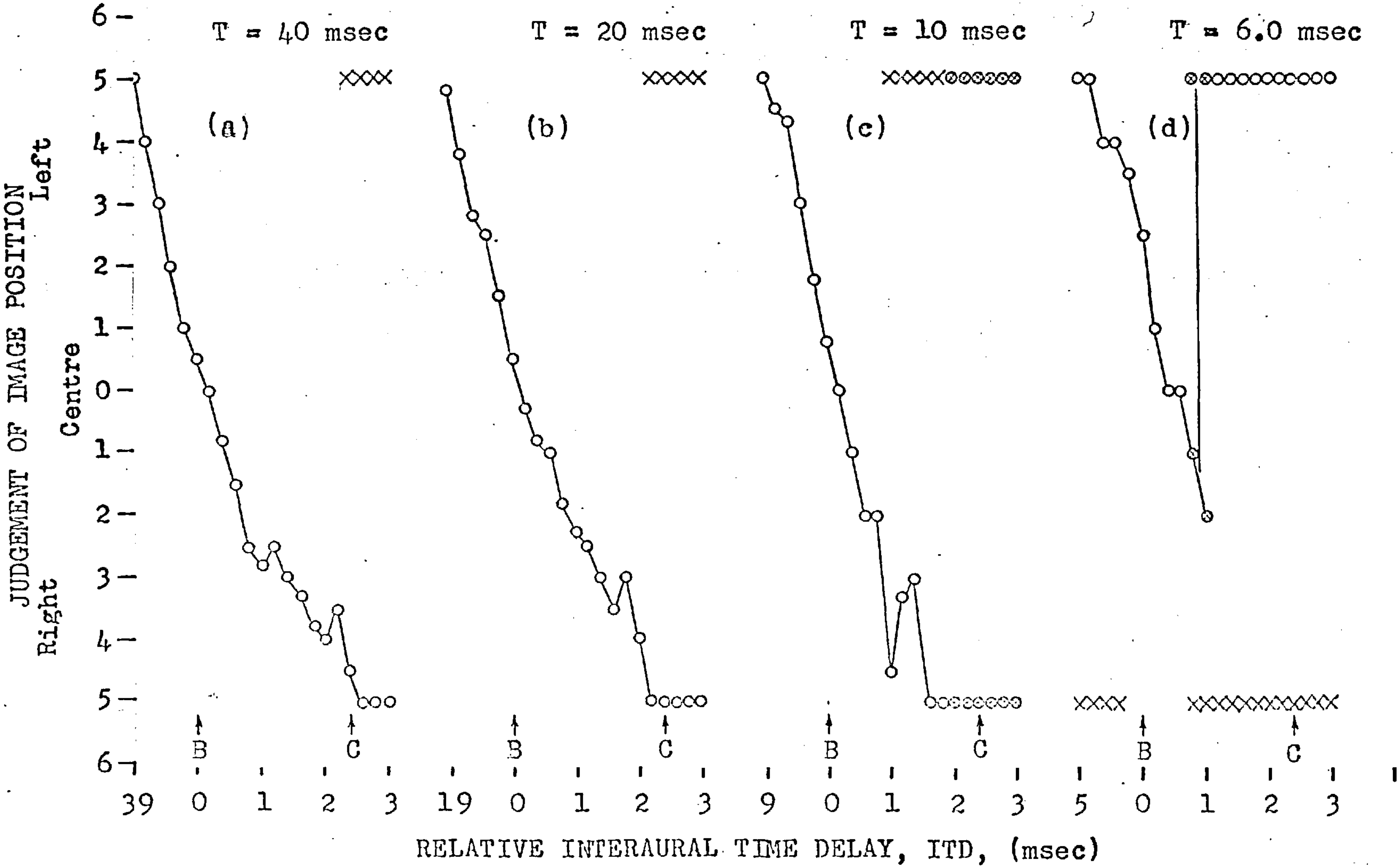


FIGURE 6.8

SUBJECT: F.E.T.
SIGNALS: T = as indicated   δt = 2.4 msec
A:B:C = 1:1:1   Unfiltered

Low-Pitched Impulsive Image

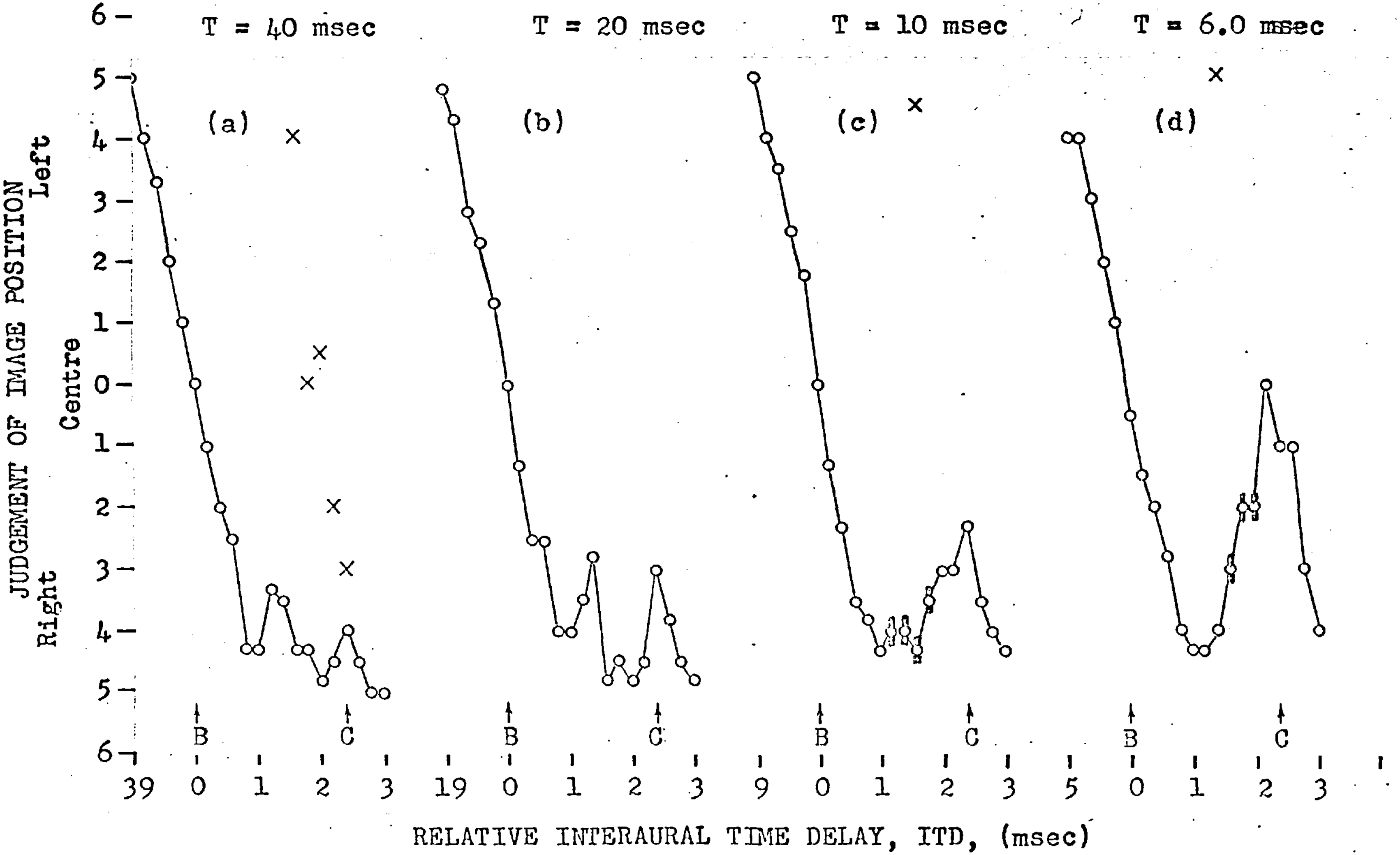


FIGURE 6.9

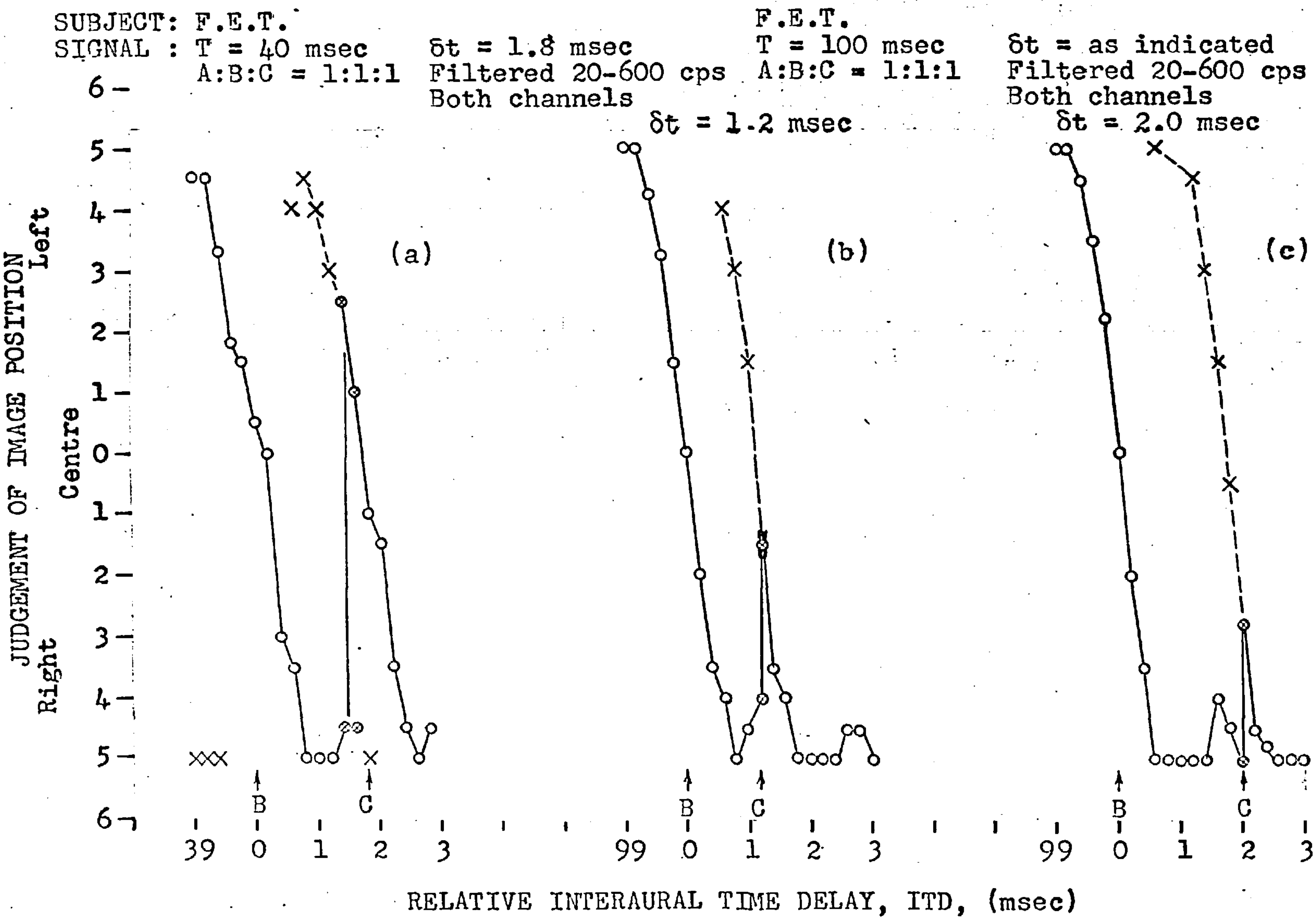
SUBJECT: F.E.T.
SIGNAL : T = 40 msec
A:B:C = 1:1:1
δt = 1.8 msec
Filtered 20-600 cps
Both channels
F.E.T.
T = 100 msec
A:B:C = 1:1:1
δt = as indicated
Filtered 20-600 cps
Both channels
δt = 1.2 msec
δt = 2.0 msec
(a)
(b)
(c)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
39
0
1
2
3
99
0
1
2
3
99
0
1
2
3
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.10

SUBJECT: F.E.T.
SIGNAL : T = 100 msec   δt = as indicated
A:B:C = 1:1:1   Filtered 6-15 kc, both channels

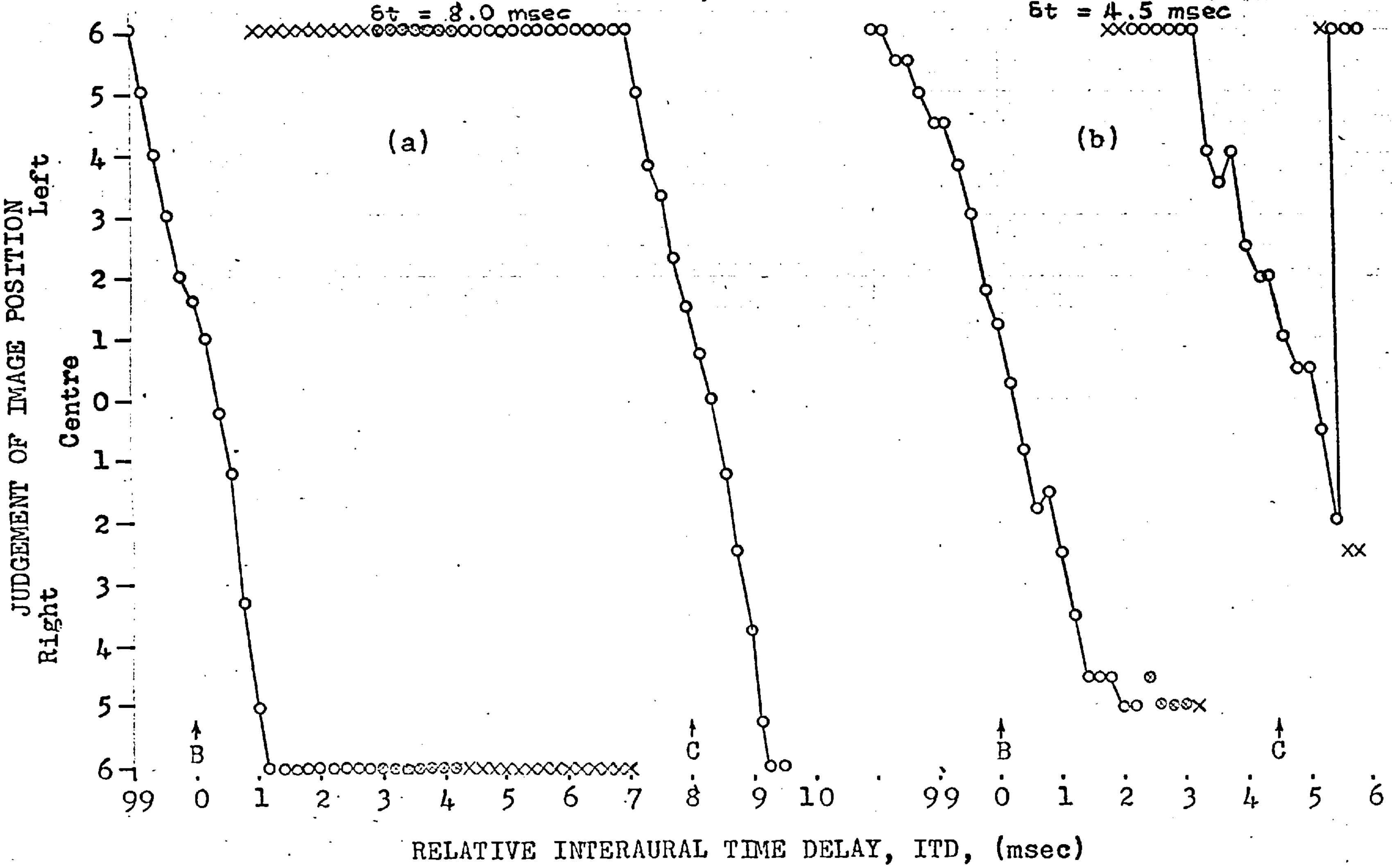


FIGURE 6.11

SUBJECT: F.E.T.
SIGNAL : T = 100 msec δt = as indicated
A:B:C = 1:1:1 Filtered 6-15 kc, both channels.

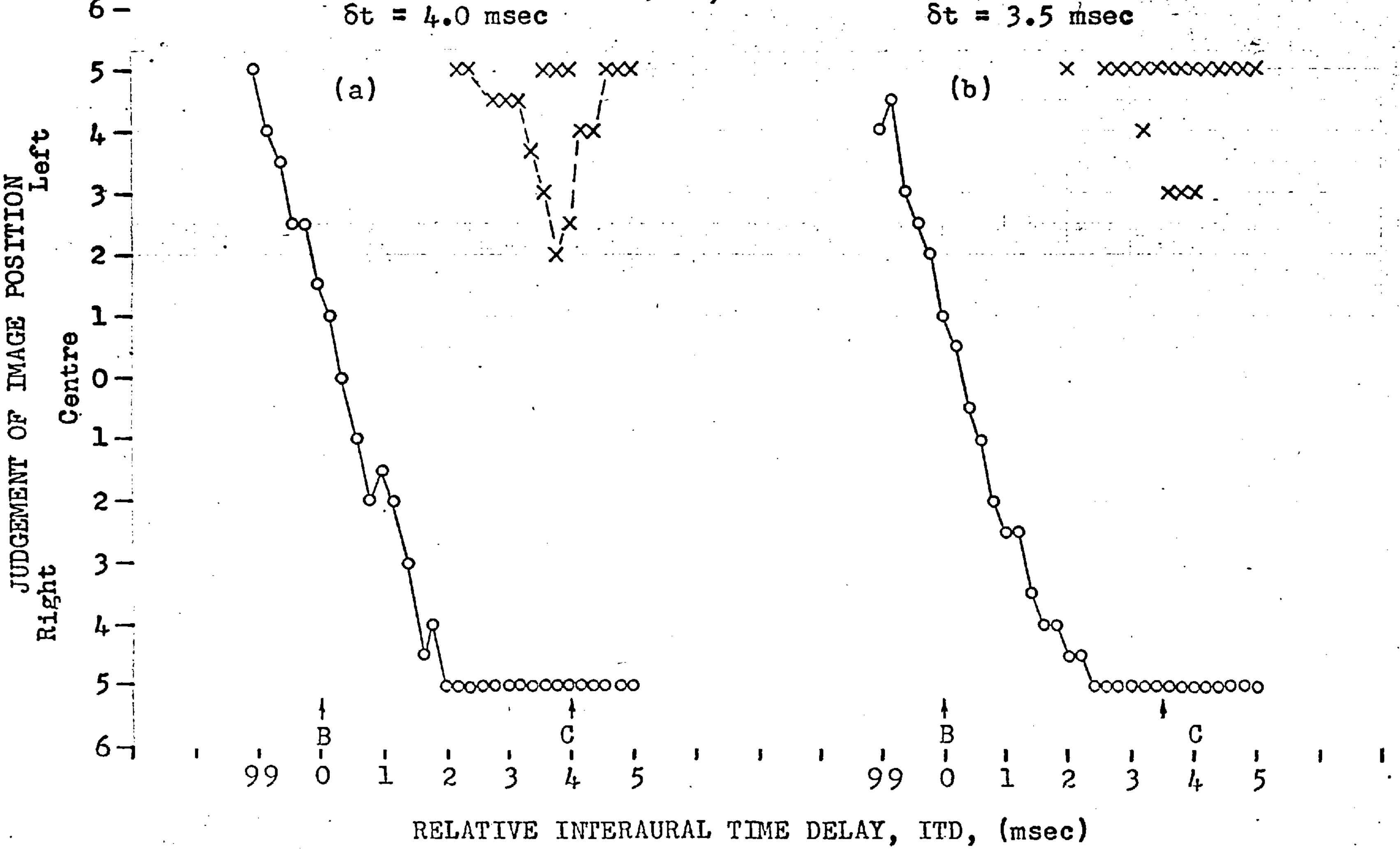
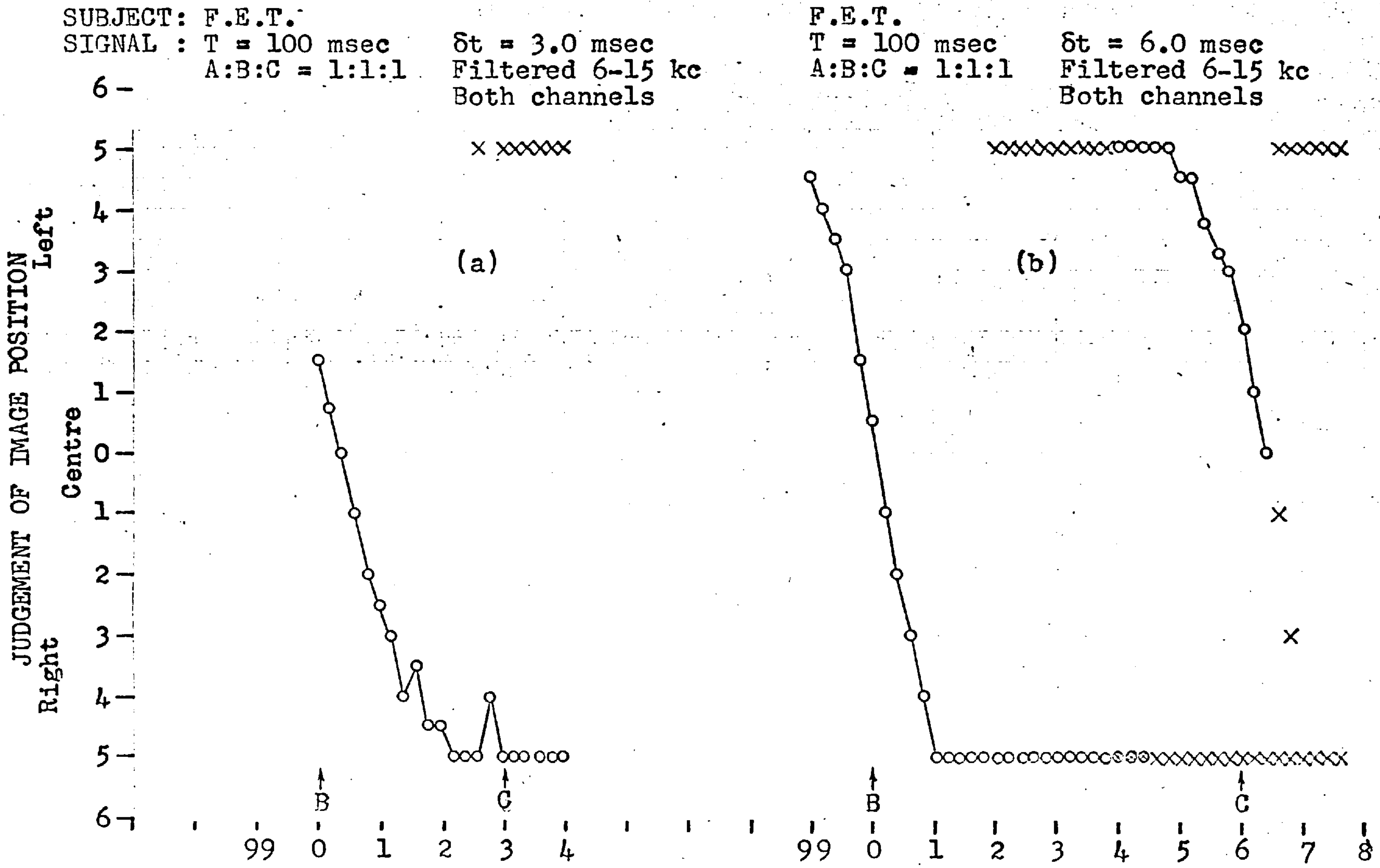


FIGURE 6.12



RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.13

SUBJECT: F.E.T.
SIGNAL : T = 100 msec δt = 6.0 msec
A:B:C = 1:1:1*   Filtered 6-15 kc, both channels.

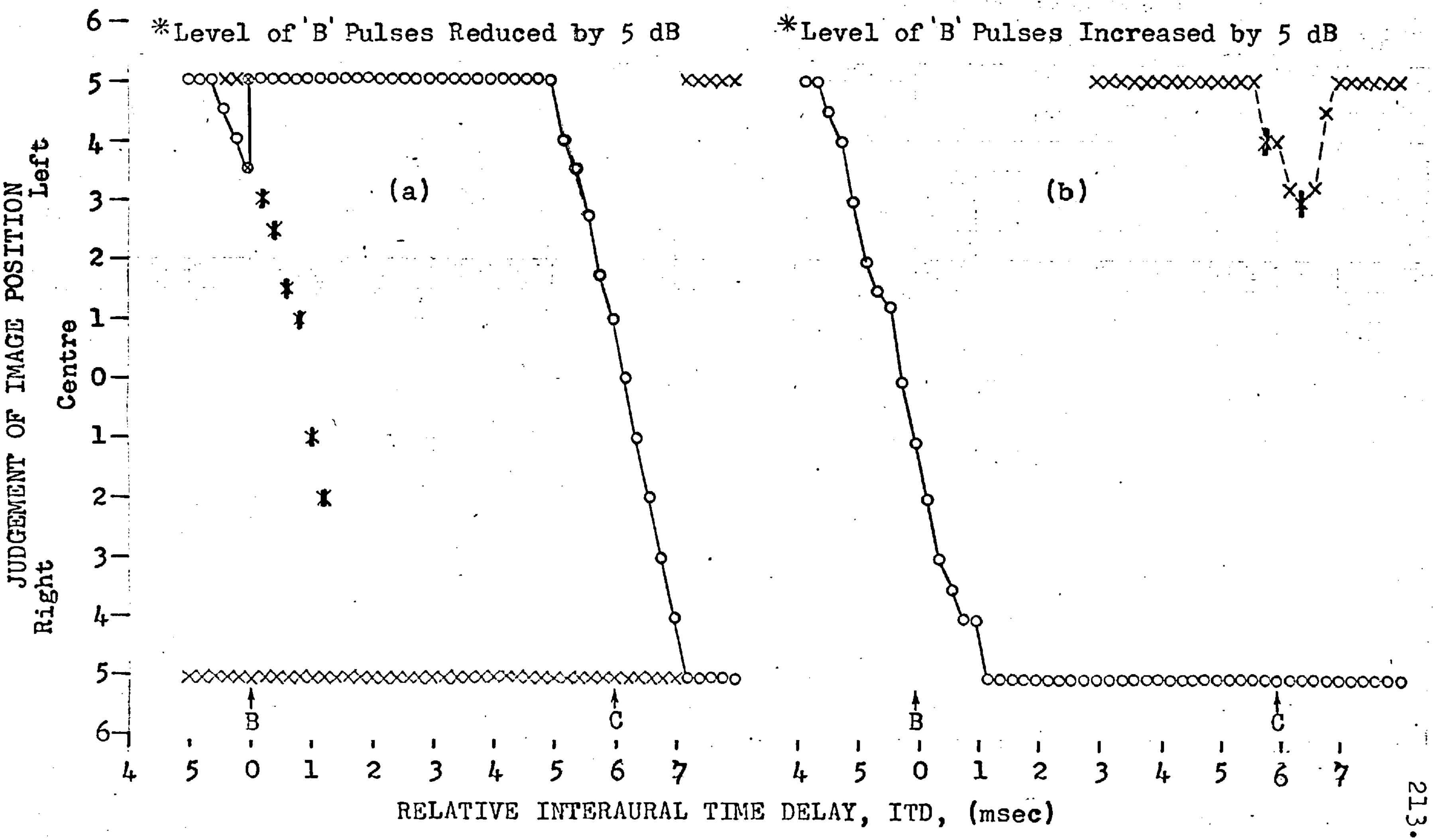


FIGURE 6.14

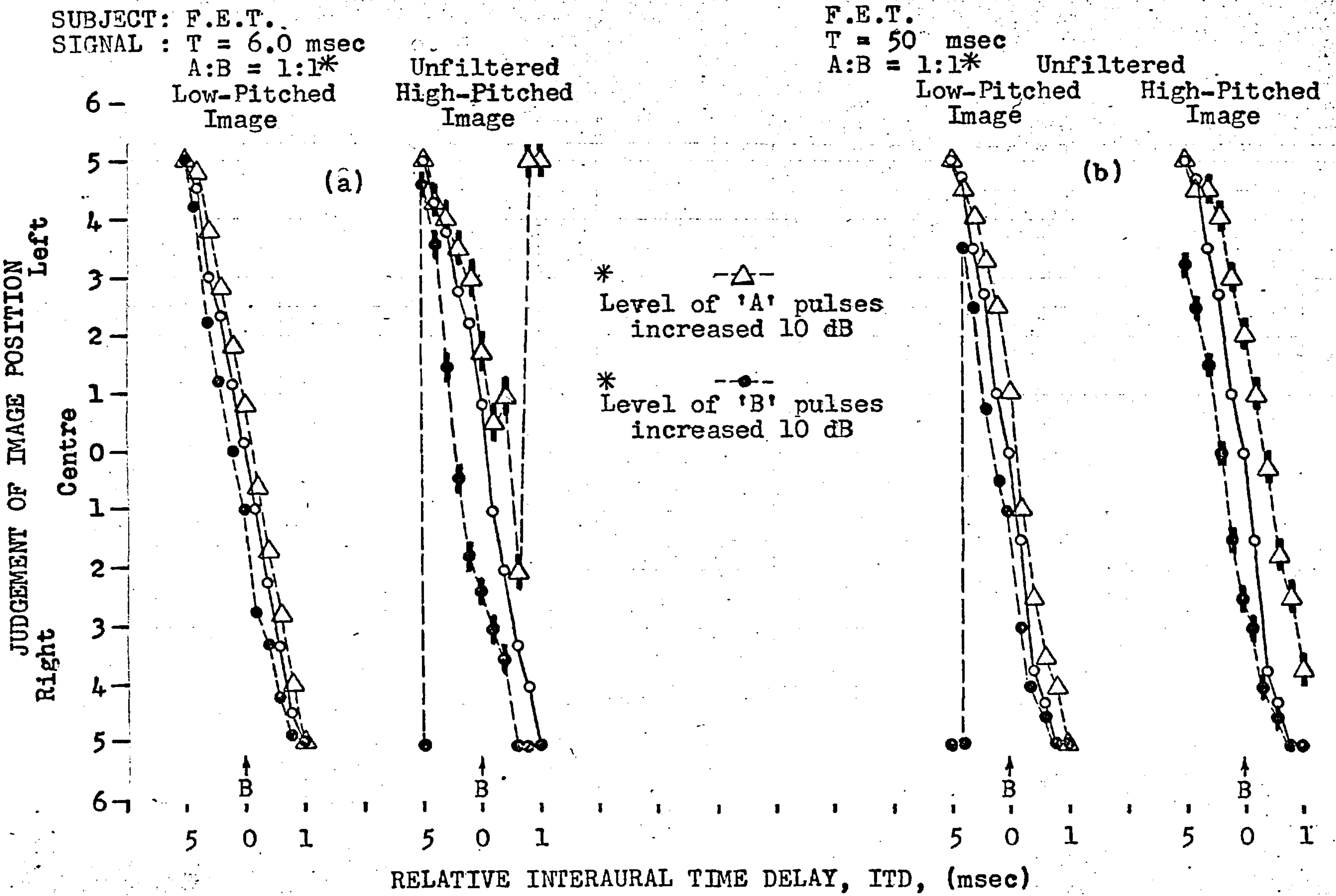

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec
A:B = 1:1*
Unfiltered
Low-Pitched Image
High-Pitched Image
(a)
F.E.T.
T = 50 msec
A:B = 1:1*
Unfiltered
Low-Pitched Image
High-Pitched Image
(b)
* –△– Level of 'A' pulses increased 10 dB
* –●– Level of 'B' pulses increased 10 dB
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.15

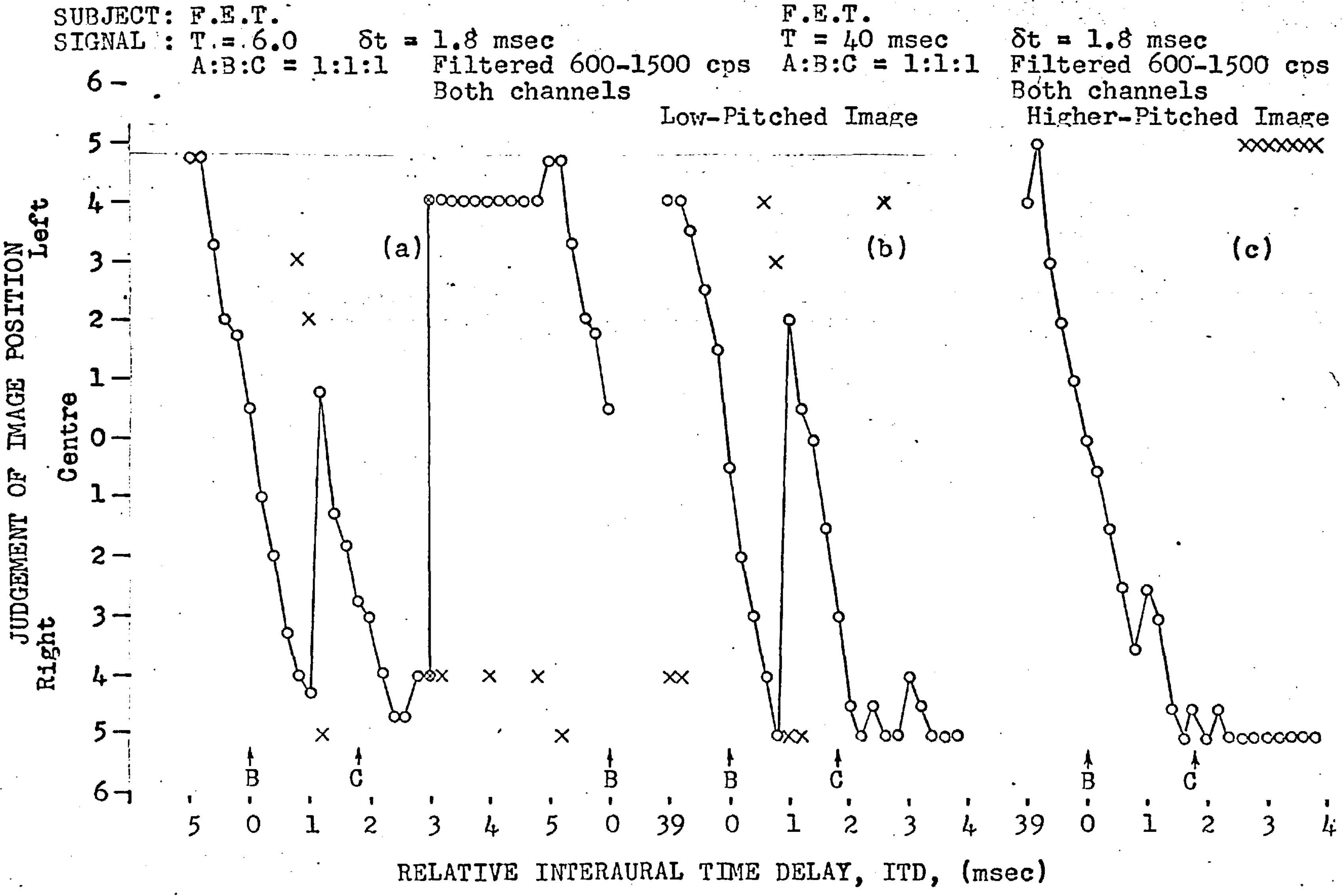
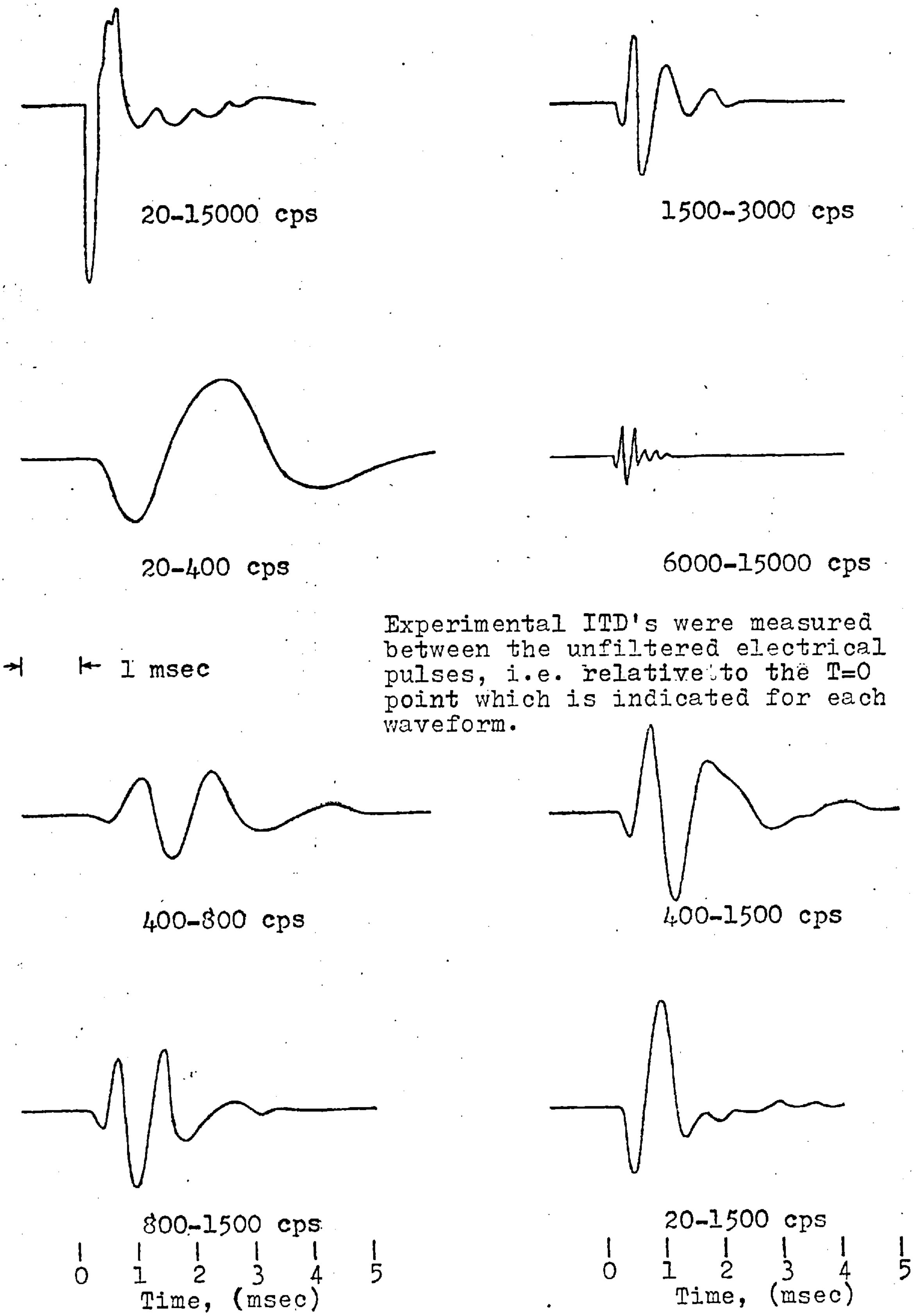

SUBJECT: F.E.T.
SIGNAL : T = 6.0 δt = 1.8 msec
A:B:C = 1:1:1 Filtered 600-1500 cps
Both channels
F.E.T.
T = 40 msec δt = 1.8 msec
A:B:C = 1:1:1 Filtered 600-1500 cps
Both channels
Low-Pitched Image
Higher-Pitched Image
(a)
(b)
(c)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
C
5
0
1
2
3
4
5
0
39
0
1
2
3
4
39
0
1
2
3
4
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.16

FIGURE 6.17

Waveforms of Filtered Acoustic Transients



20-15000 cps

1500-3000 cps

20-400 cps

6000-15000 cps

Experimental ITD's were measured between the unfiltered electrical pulses, i.e. relative to the T=0 point which is indicated for each waveform.

400-800 cps

400-1500 cps

800-1500 cps

20-1500 cps

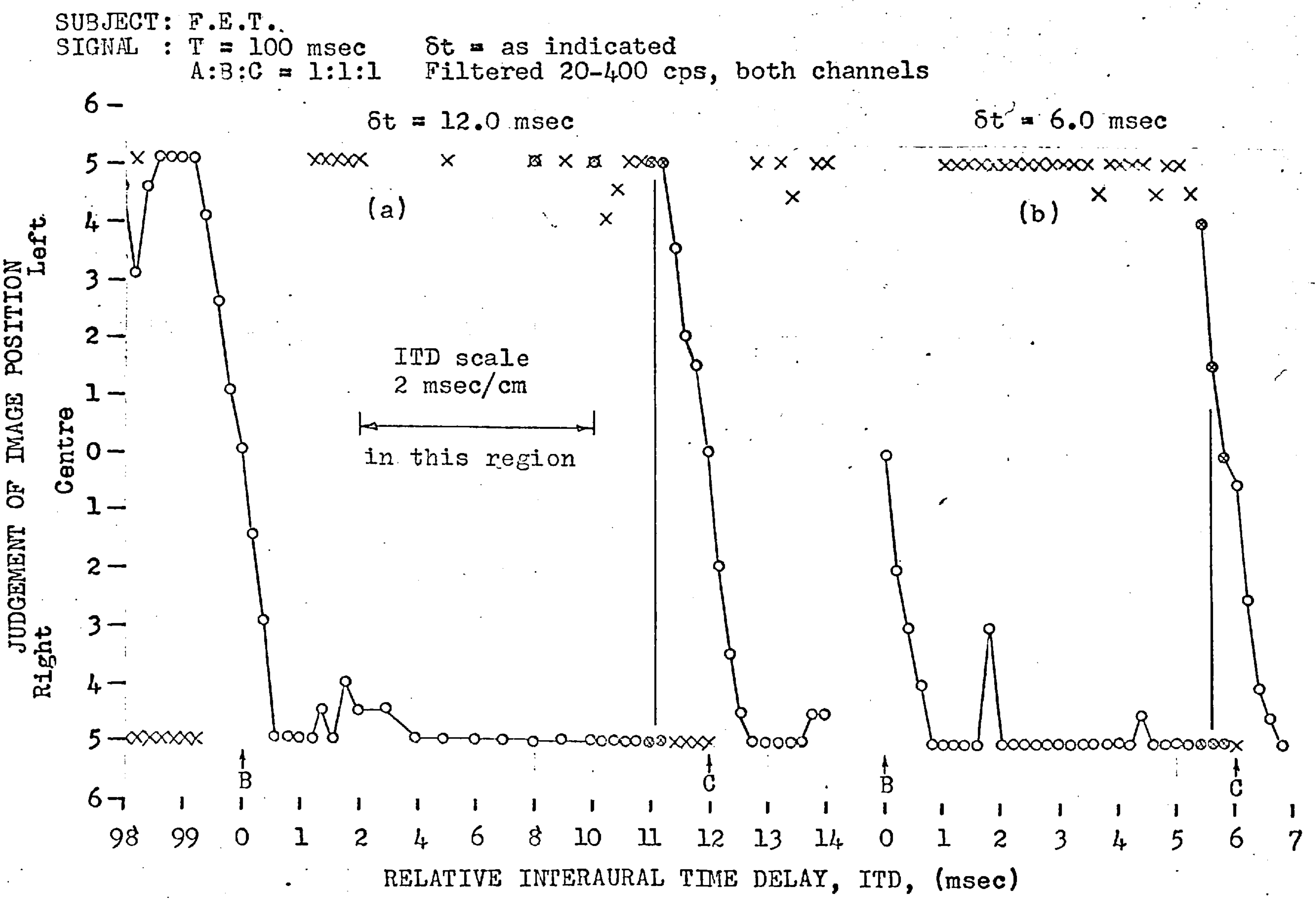

SUBJECT: F.E.T.
SIGNAL : T = 100 msec δt = as indicated
A:B:C = 1:1:1 Filtered 20-400 cps, both channels
δt = 12.0 msec
δt = 6.0 msec
(a)
(b)
ITD scale
2 msec/cm
in this region
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
C
98
99
0
1
2
4
6
8
10
11
12
13
14
0
1
2
3
4
5
6
7
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.18

SUBJECT: F.E.T.
SIGNAL : T = 100 msec   δt = as indicated
A:B:C = 1:1:1   Filtered 20-400 cps, both channels

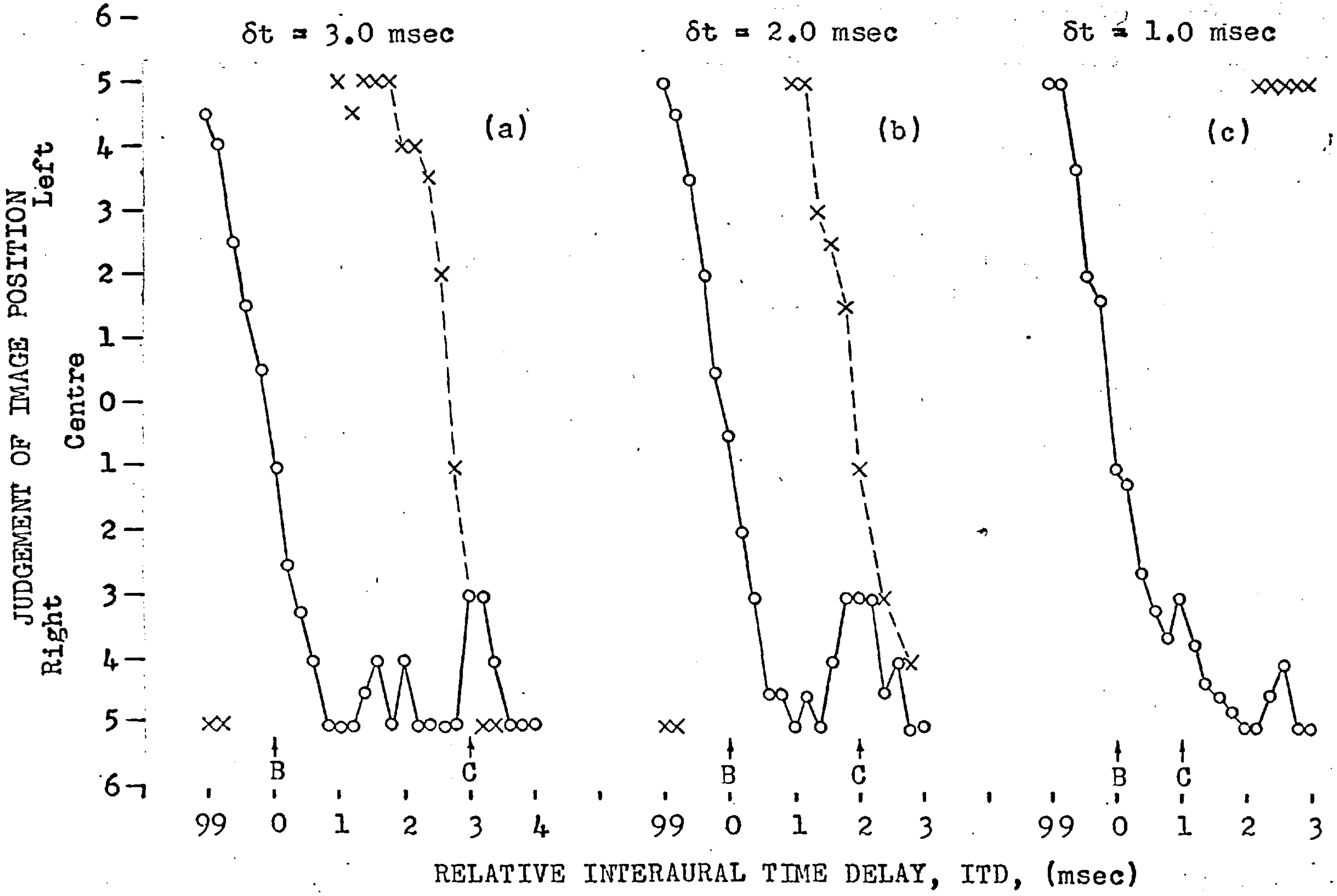


FIGURE 6.19

SUBJECT: F.E.T..
SIGNAL : T = 100 msec   δt = as indicated
A:B:C = 1:1:1   Filtered 400-800 cps, both channels

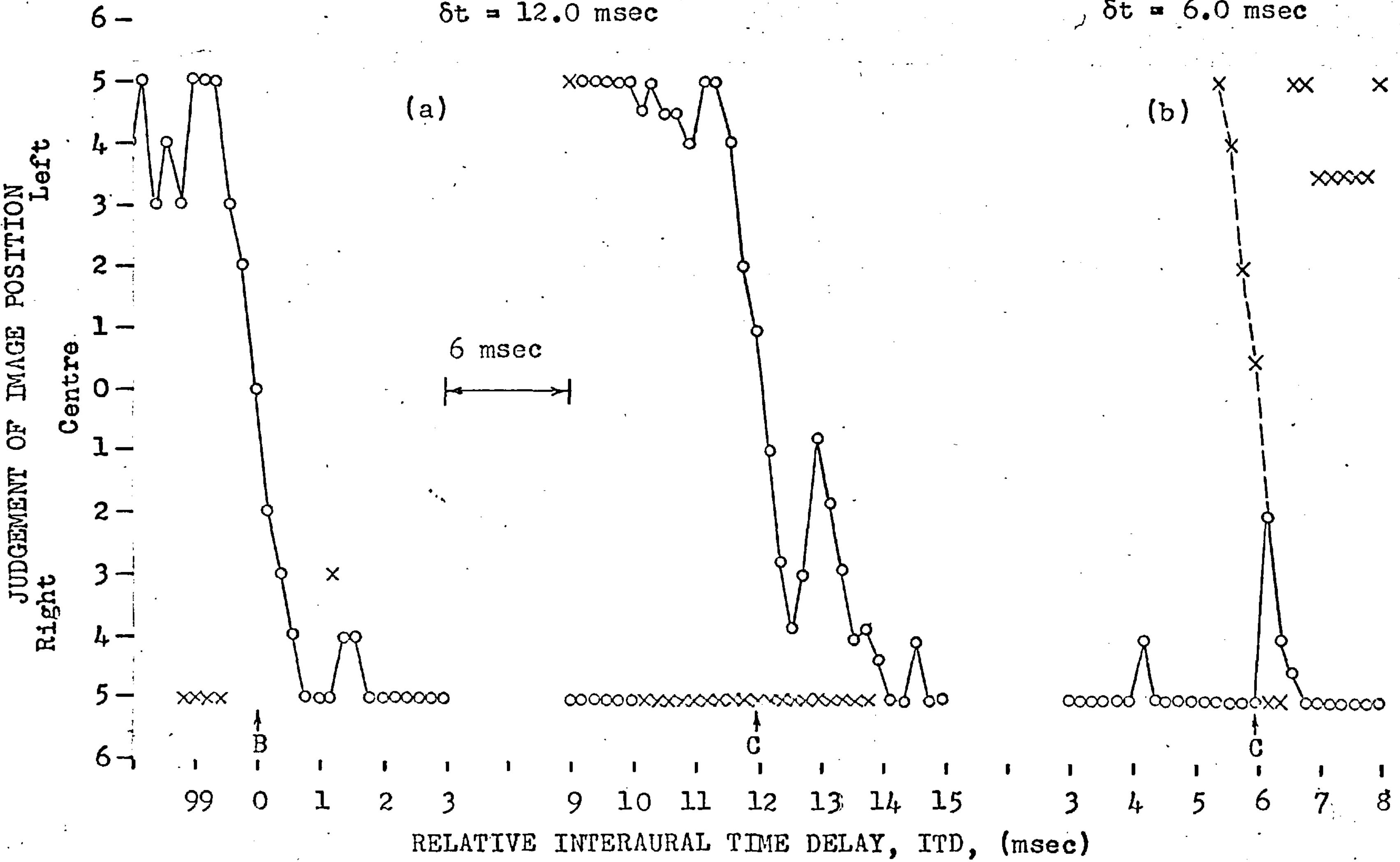


FIGURE 6.20

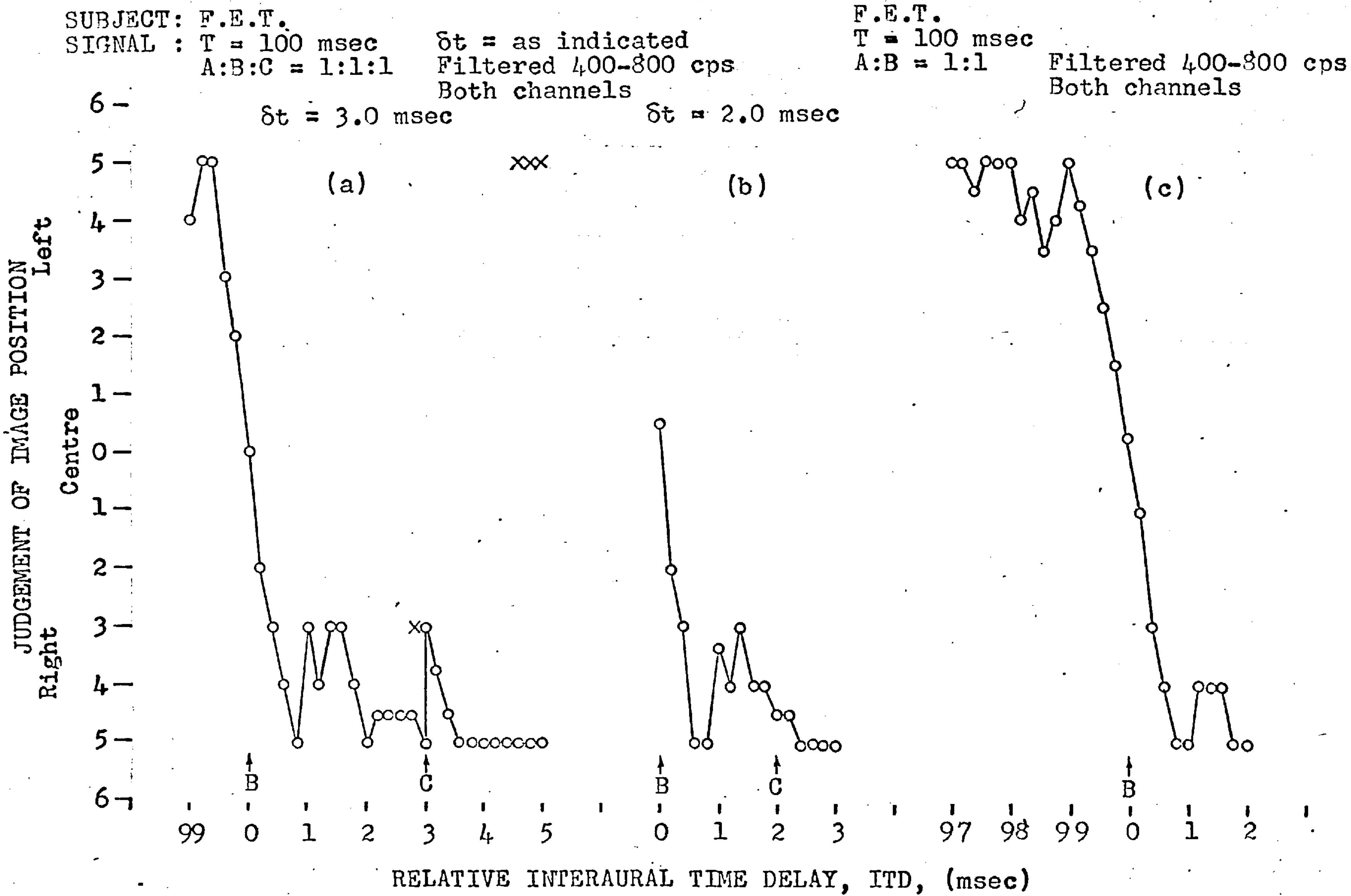

SUBJECT: F.E.T.
SIGNAL : T = 100 msec
A:B:C = 1:1:1
δt = as indicated
Filtered 400-800 cps
Both channels
δt = 3.0 msec
δt = 2.0 msec
F.E.T.
T = 100 msec
A:B = 1:1
Filtered 400-800 cps
Both channels
(a)
(b)
(c)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
C
B
99
0
1
2
3
4
5
0
1
2
3
97
98
99
0
1
2
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.21

SUBJECT: F.E.T.
SIGNAL : T = 100 msec δt = as indicated
A:B:C = 1:1:1 Filtered 800-1500 cps
Both channels

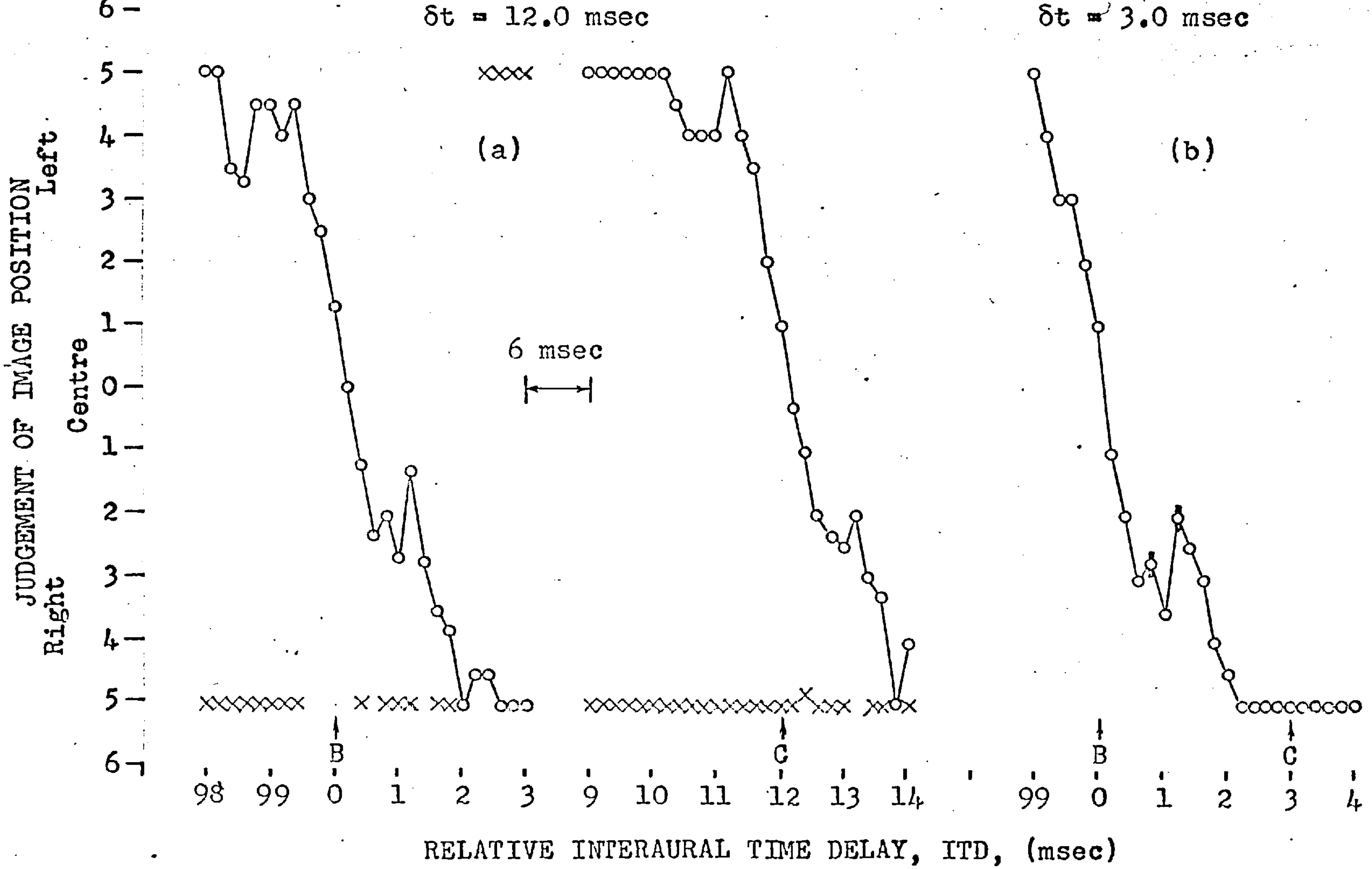


FIGURE 6.22

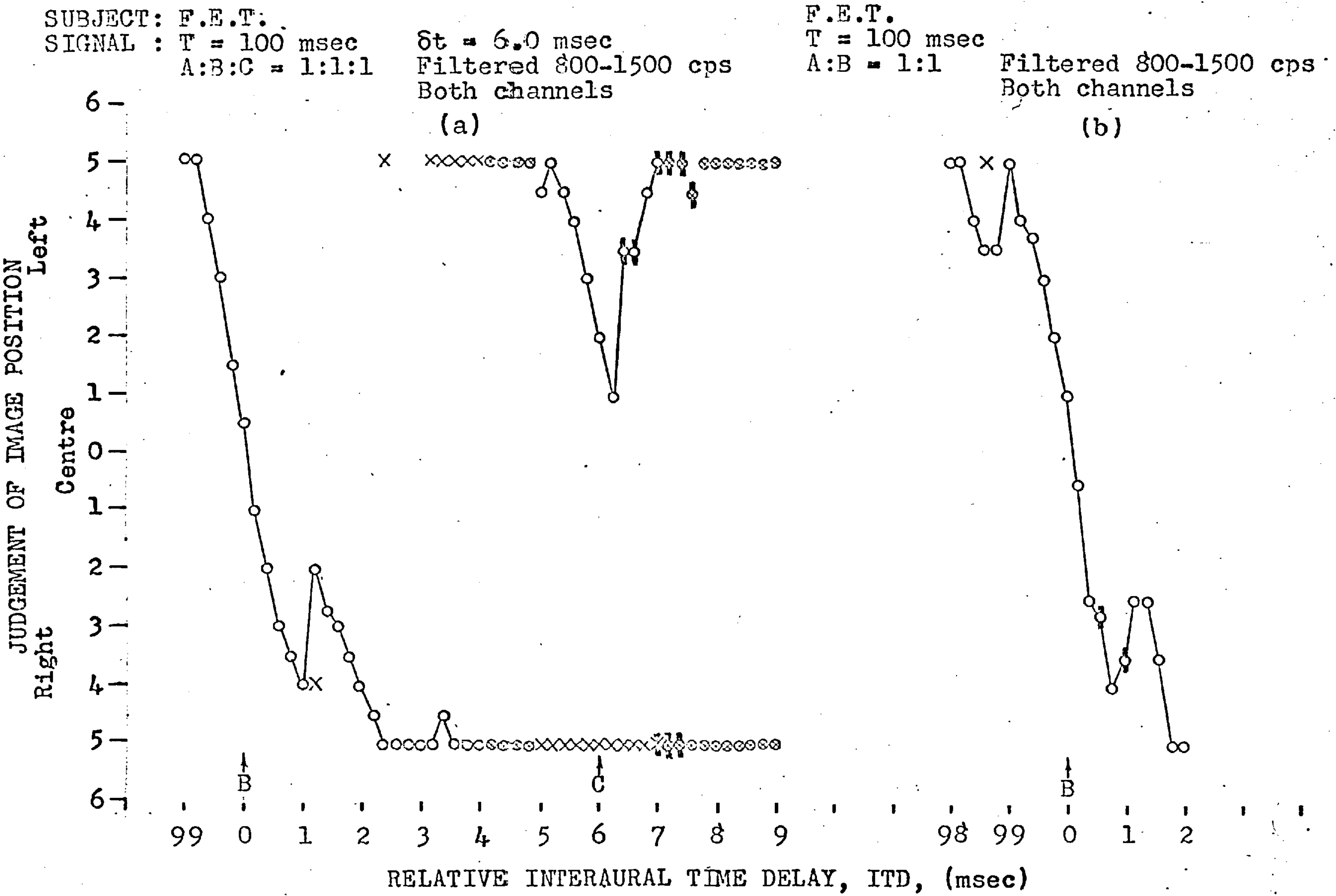

SUBJECT: F.E.T.
SIGNAL : T = 100 msec
A:B:C = 1:1:1
δt = 6.0 msec
Filtered 800-1500 cps
Both channels
(a)
F.E.T.
T = 100 msec
A:B = 1:1
Filtered 800-1500 cps
Both channels
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
99
0
1
2
3
4
5
6
7
8
9
98
99
0
1
2
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.23

SUBJECT: F.E.T.
SIGNAL : T = 100 msec δt = as indicated
A:B:C = 1:1:1 Filtered 1500-3000 cps, both channels

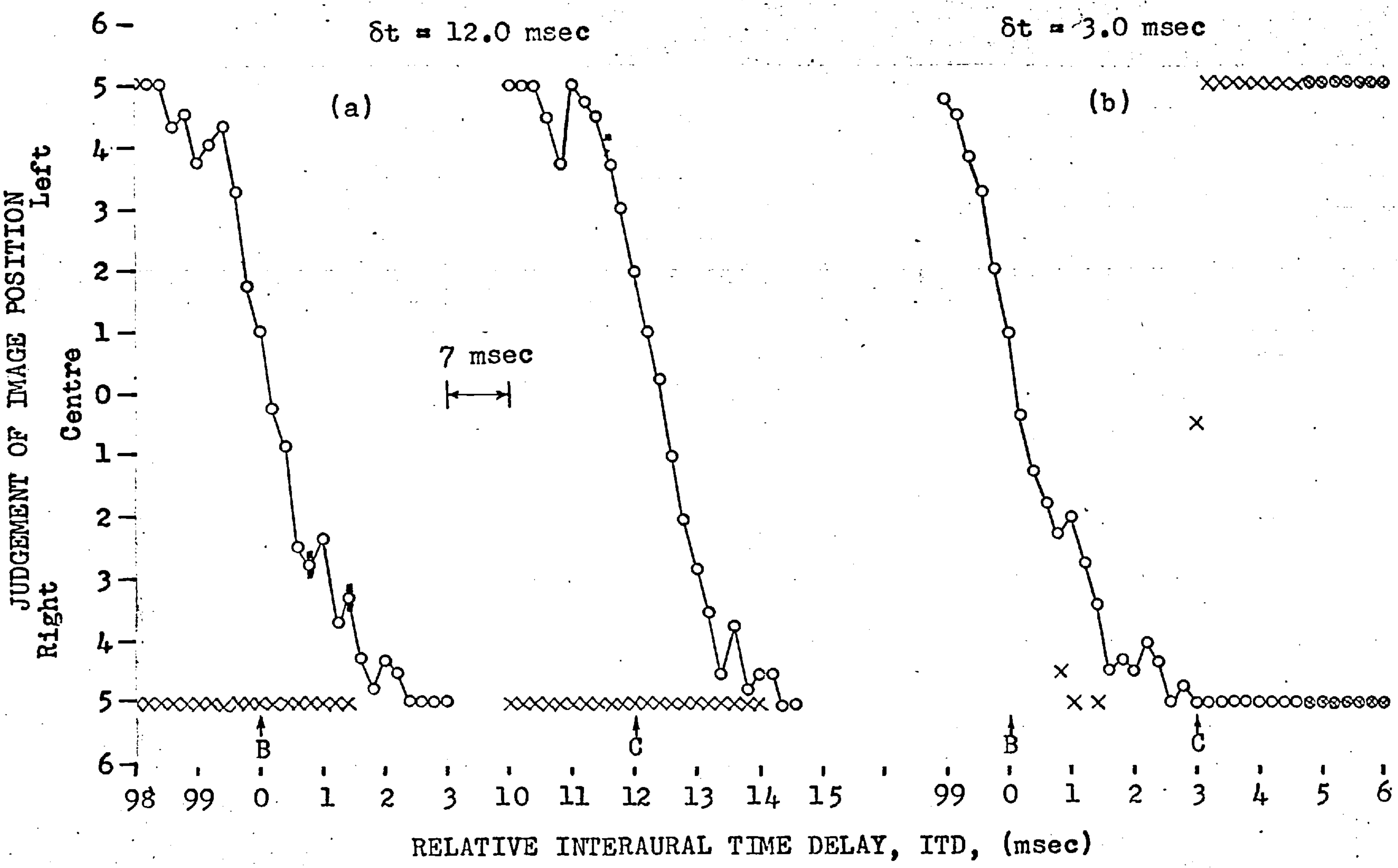


FIGURE 6.24

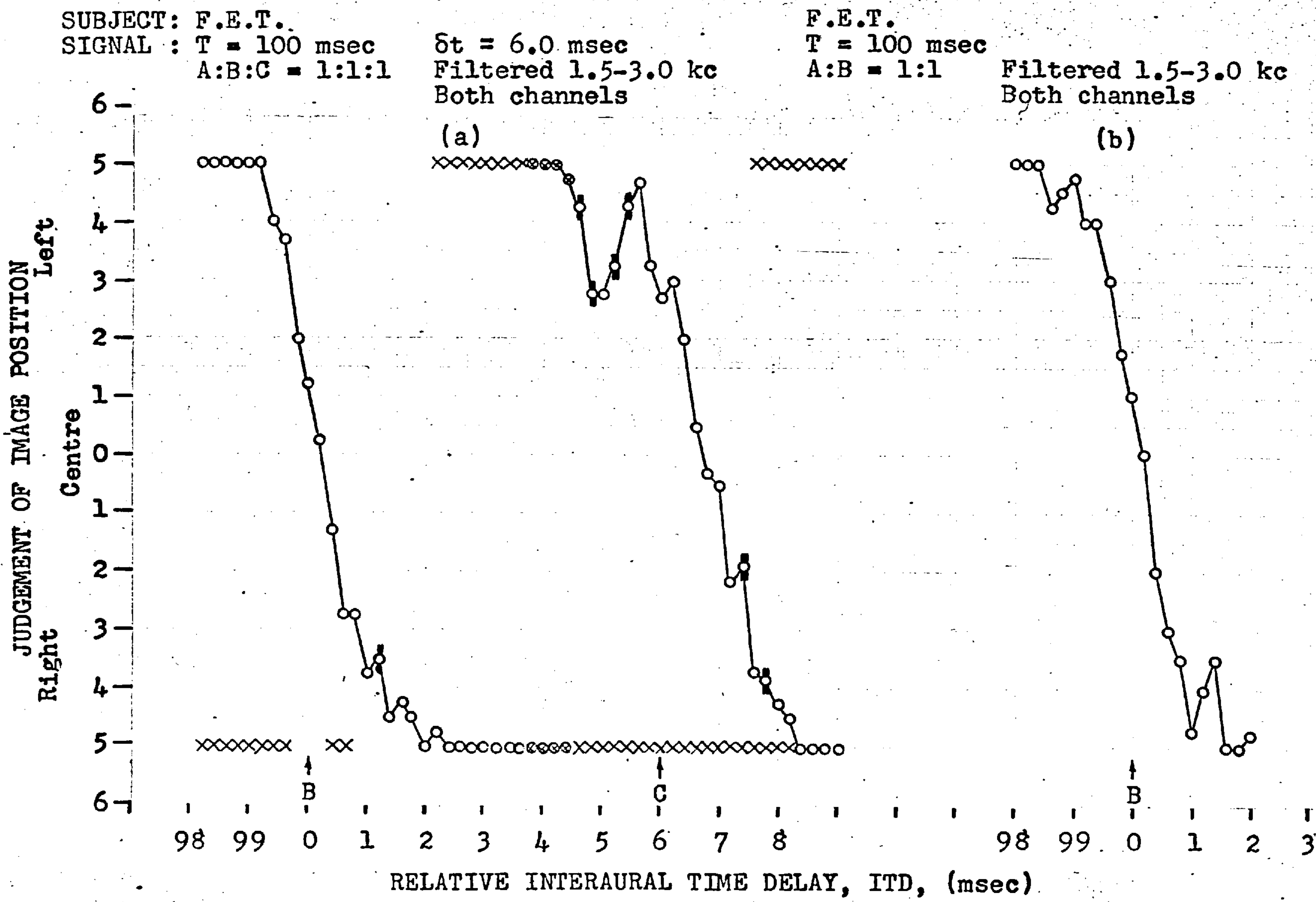

SUBJECT: F.E.T.
SIGNAL : T = 100 msec
A:B:C = 1:1:1
δt = 6.0 msec
Filtered 1.5-3.0 kc
Both channels
(a)
F.E.T.
T = 100 msec
A:B = 1:1
Filtered 1.5-3.0 kc
Both channels
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
98
99
0
1
2
3
4
5
6
7
8
98
99
0
1
2
3
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.25

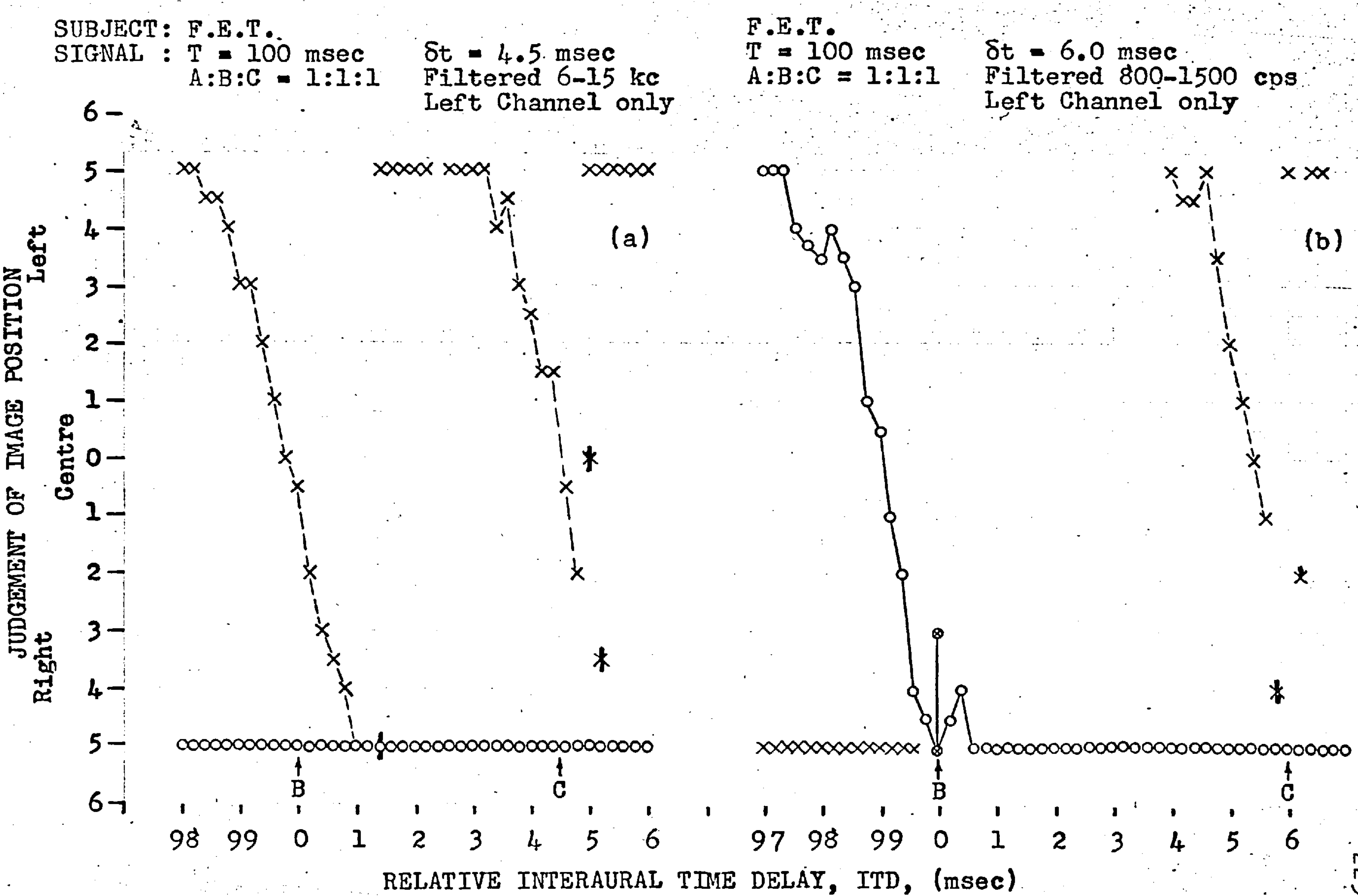

SUBJECT: F.E.T.
SIGNAL : T = 100 msec
A:B:C = 1:1:1
δt = 4.5 msec
Filtered 6-15 kc
Left Channel only
F.E.T.
T = 100 msec
A:B:C = 1:1:1
δt = 6.0 msec
Filtered 800-1500 cps
Left Channel only
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
C
98
99
0
1
2
3
4
5
6
97
98
99
0
1
2
3
4
5
6
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.26

②

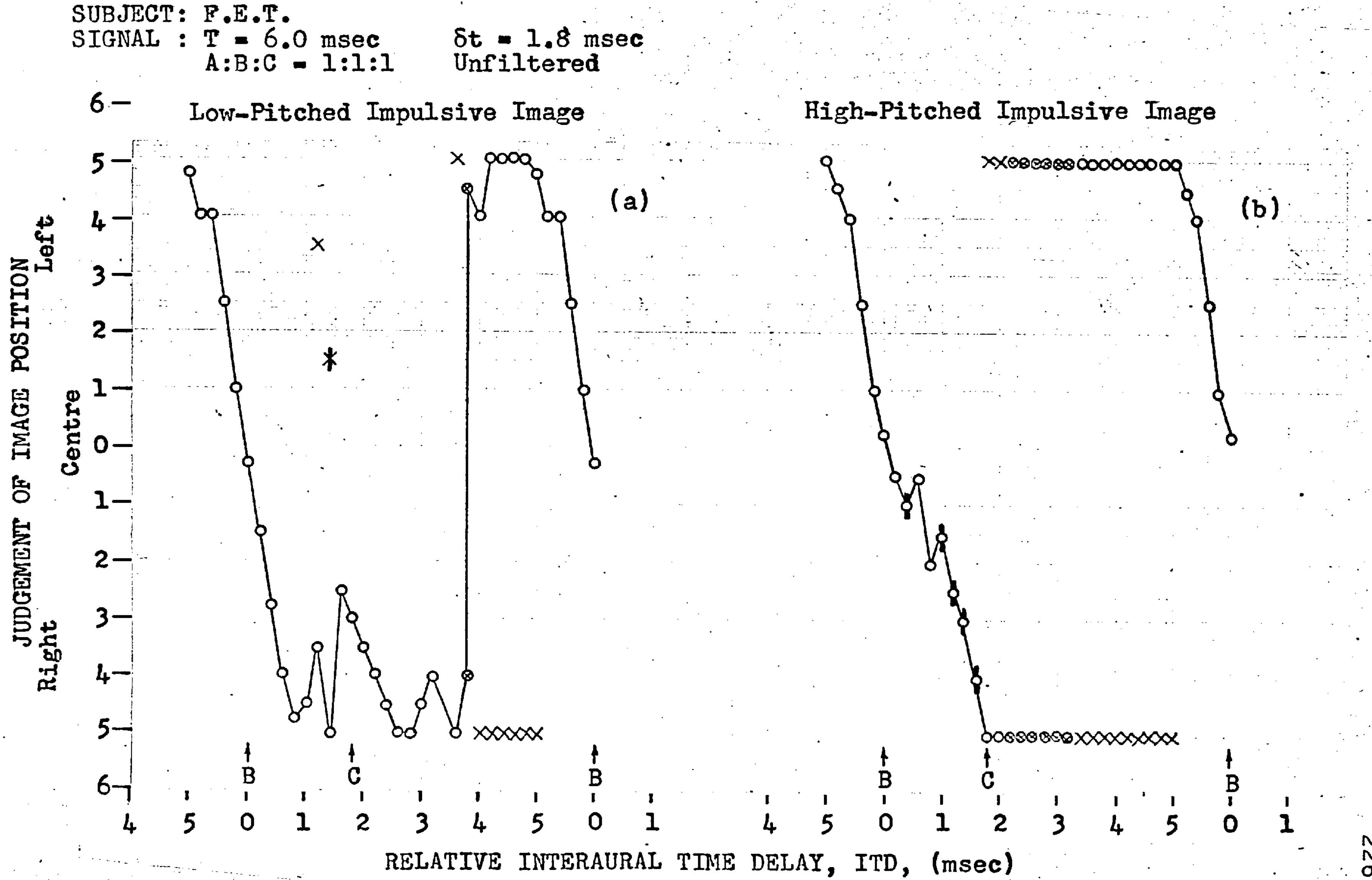

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec   δt = 1.8 msec
A:B:C = 1:1:1   Unfiltered
Low-Pitched Impulsive Image
High-Pitched Impulsive Image
(a)
(b)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
C
B
B
C
B
4
5
0
1
2
3
4
5
0
1
4
5
0
1
2
3
4
5
0
1
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 6.27

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec   δt = 1.8 msec
A:B:C = 1:1:1   Unfiltered
MASKING NOISE: Right Channel, filtered as follows:

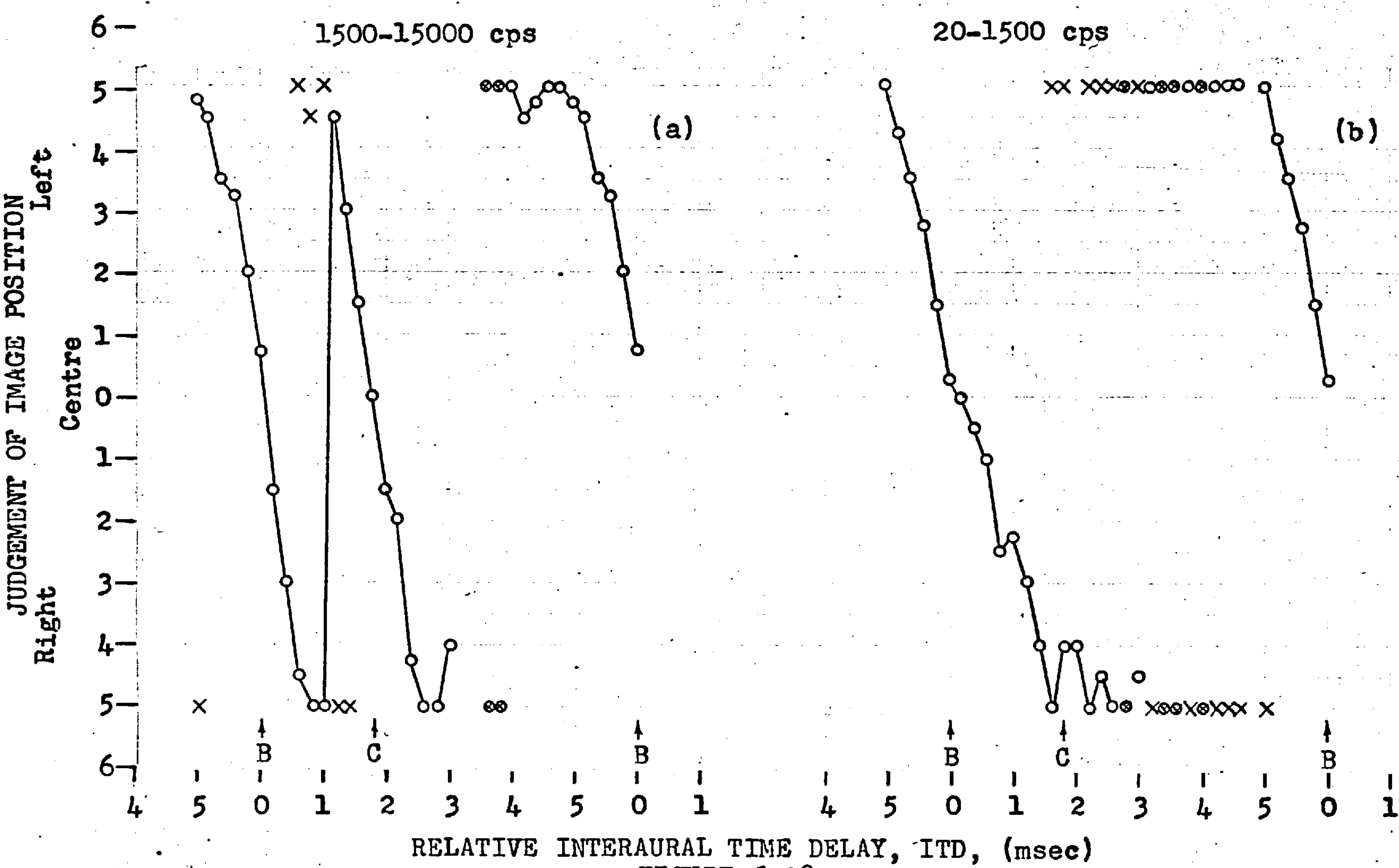


FIGURE 6.28

②

SUBJECT: F.E.T.
SIGNALS: T = 6.0 msec   δt = 1.8 msec
A:B:C = 1:1:1   Unfiltered
MASKING NOISE: Right Channel, filtered 6000-15000 cps

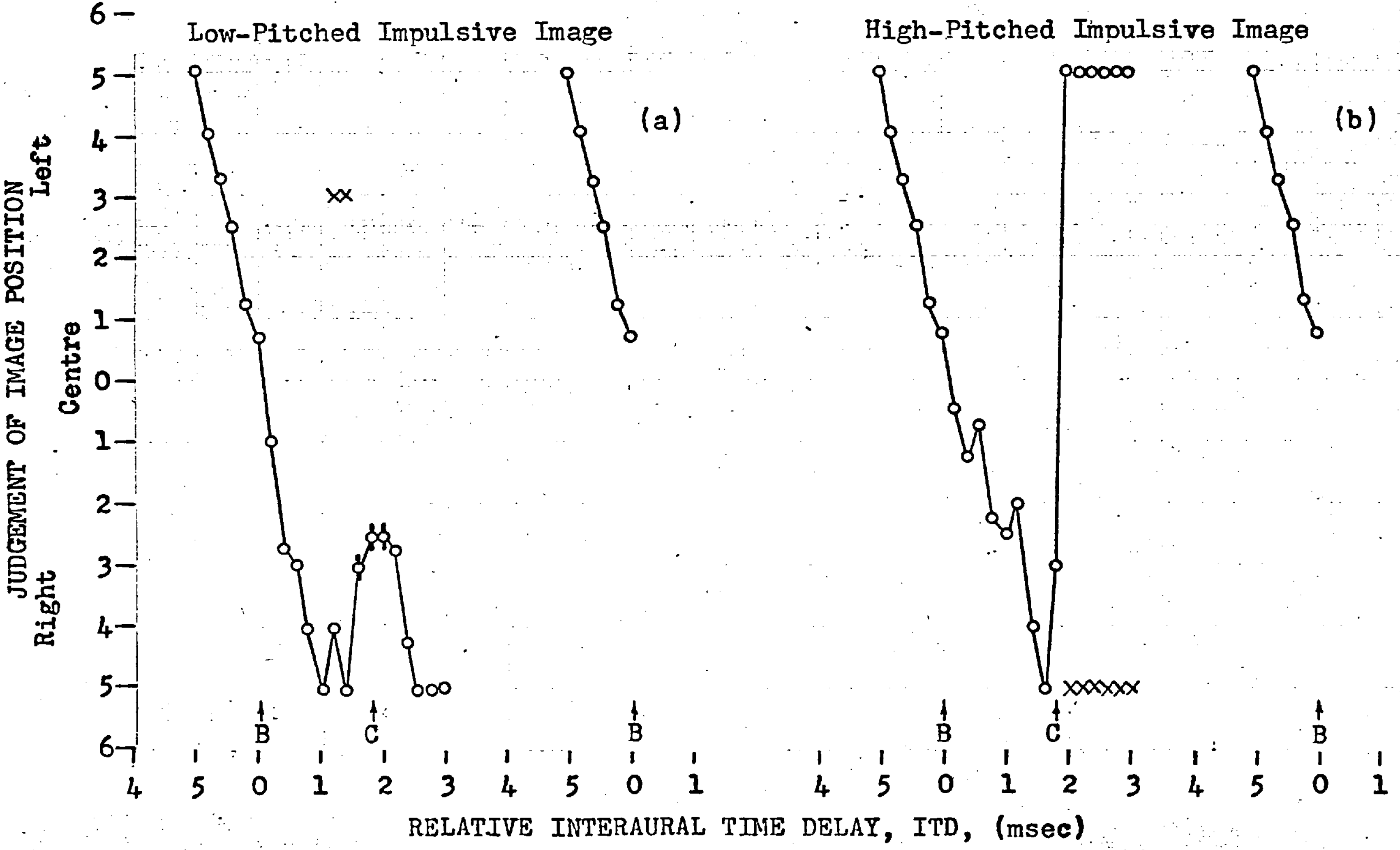


FIGURE 6.29

SUBJECT: F.E.T.
SIGNAL : T = 20 msec
A:B = 1:1 Unfiltered
MASKING NOISE: Right Channel, filtered as follows:

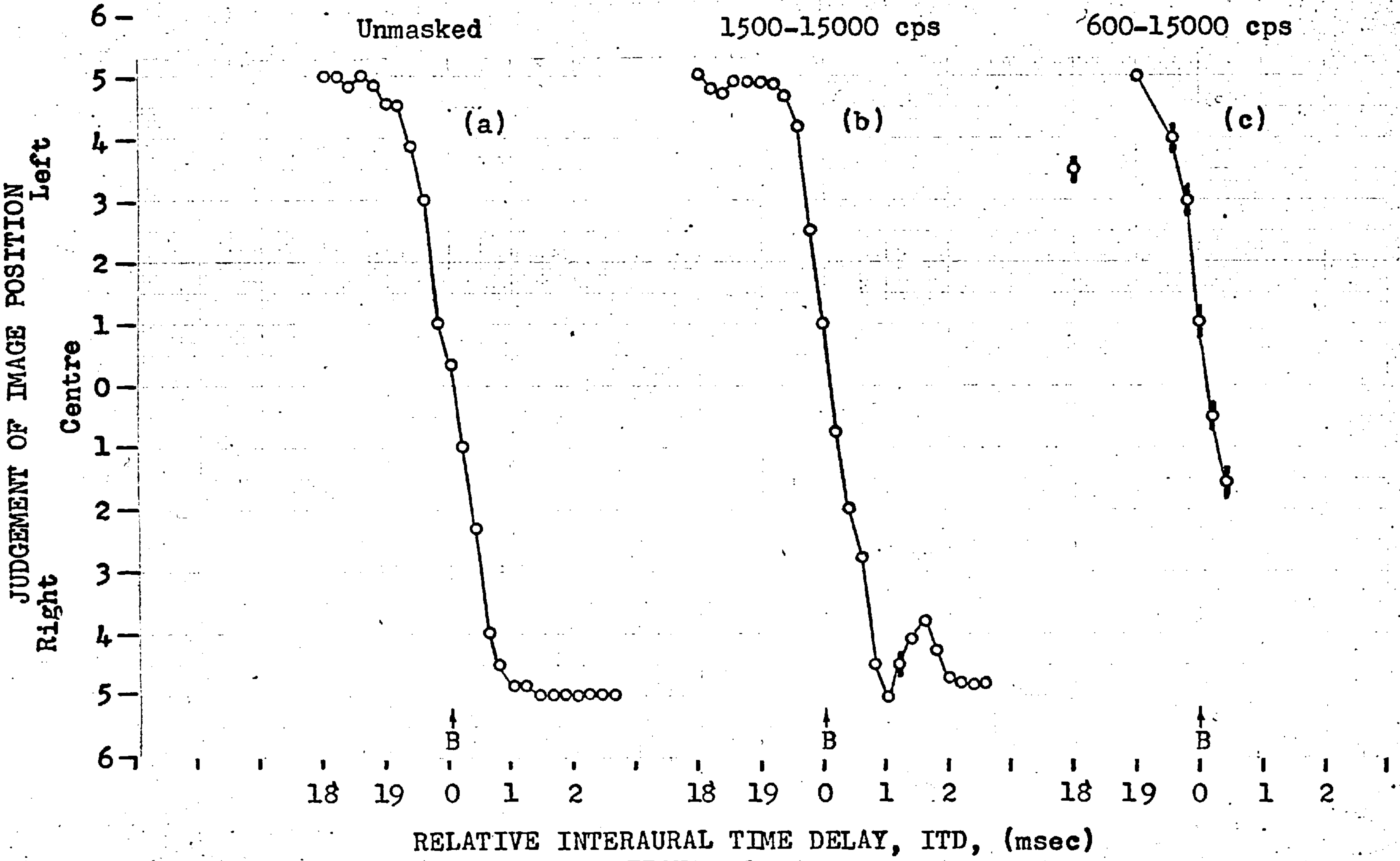


FIGURE 6.30

SUBJECT: F.E.T.
SIGNAL : T = 6.0 msec   Filtered: 20-400 cps, Left Channel
         A:B = 2:1
MASKING NOISE: Right Channel, Filtered as follows:

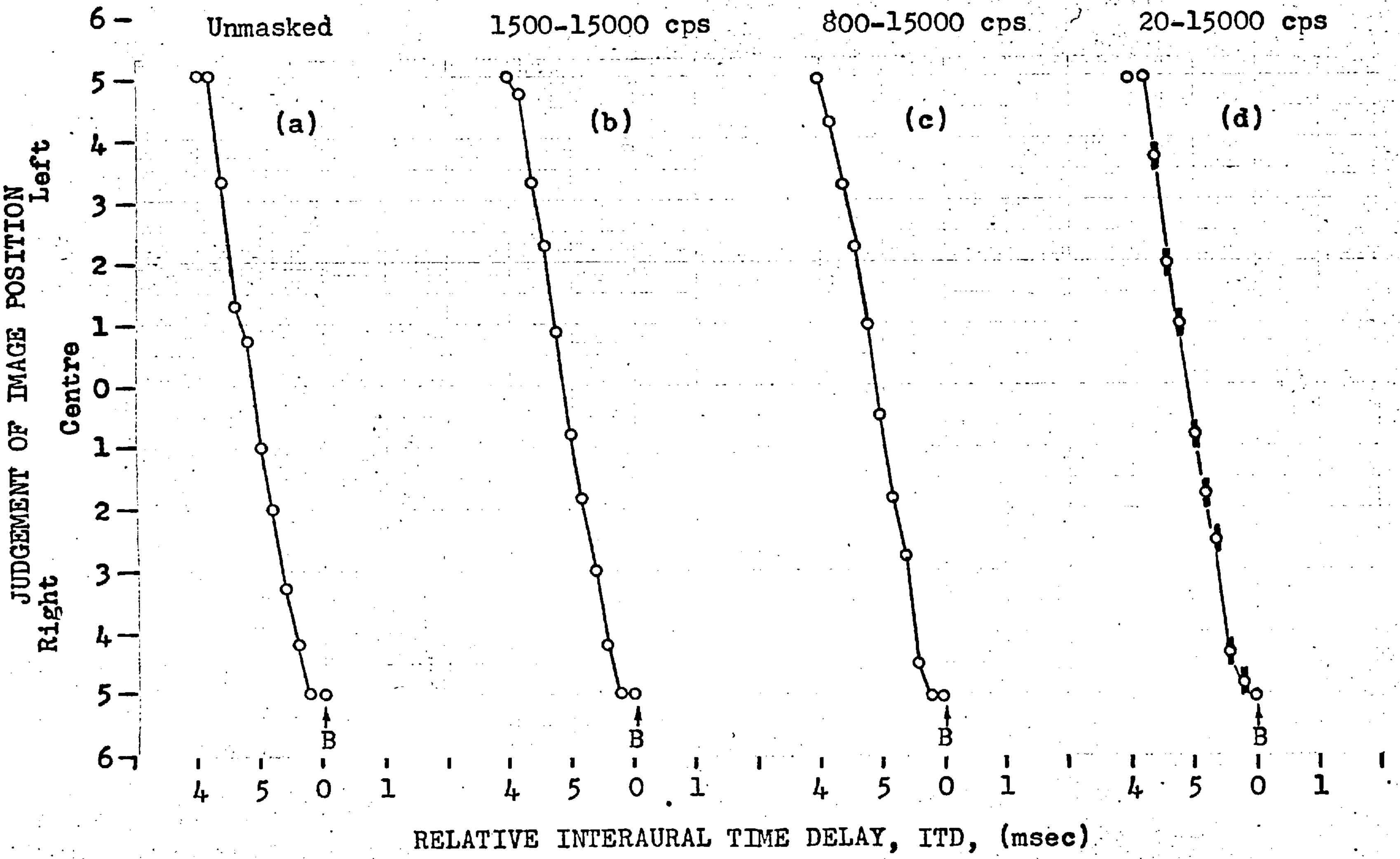


FIGURE 6.31

CHAPTER 7

THE PERIPHERAL AUDITORY MECHANISM AND ITS RELATION TO BINAURAL ACOUSTIC IMAGES

7.1.0 Introduction

The preceding investigations of acoustic image phenomena have been instructive in a number of respects, but there were a number of points which could not be clearly understood without some further consideration of auditory neurophysiology. The following investigations were intended, therefore, to cover those aspects of the subject which appear to be important in interpreting certain experimental observations reported here. Such investigations are, of course, precursors of a more comprehensive study necessary for a complete understanding of binaurally created acoustic images.

7.2.0 Experimental Technique

Experiments discussed in the previous chapter pointed out that acoustic image lateralizations were apparently related in an important sense, to the dynamic properties of the cochlea. In order to test the validity of the proposition and to enable the development of a more complete view of the auditory mechanism, it was crucial that, inter alia, the characteristics of basilar membrane displacements be generally defined. In particular, it would be desirable to know how the basilar membrane responds to acoustic transient signals. Direct evidence of this type is extremely limited, and one must rely largely on data acquired using mechanical, electrical, or computational models of the basilar membrane; even so, published data is limited. It was therefore decided that the purposes of the present study would be conveniently served by the use of such a model.

The chosen model of the basilar membrane was that developed by Flanagan (1960, 1962), which was based on the direct measurements of von Békésy. The model was constructed in its electrical network form (Flanagan, 1962), but was slightly simplified.

The low-pitched impulsive image phenomena which were of immediate interest, were clearly related to neural activity arising from the apical region of the cochlea (e.g., see section 6.4.5), i.e. to those regions maximally responsive to frequencies below about 1500 cps. For this reason the network simulating the transmission properties of the middle ear was omitted as it only affects to any important extent, the spectrum above about 1500 cps. Omitted also was the delay line, the effect of which could be easily computed and accommodated in the result. In fact, this parameter was of little concern as it turned out that important information was derived from networks representing similar distance co-ordinates of left-and right-ear basilar membranes. These would have closely similar delays.

The purpose of the model in this study was to provide an indication of the form of basilar membrane displacement at various points when the 'ear' was stimulated by various acoustic transient waveforms. This was accomplished by applying the earphone to the artificial ear, and the voltage output of the artificial ear to the input of the model. The output of the model was displayed on an oscilloscope, and could, if desired, be photographed using standard procedures. In this way the model was stimulated by a signal which was, presumably, a reasonable approximation to the actual acoustic output of the earphone.

Acoustic-image lateralization experiments using various acoustic transient signals were conducted along the same lines as those of previous chapters. These results were used in conjunction with corresponding observations of basilar membrane displacement on the model to seek explanations of certain acoustic-image lateralization phenomena.

A few experiments were conducted using a centering technique. These experiments were thus arranged to enable rapid assessment of a number of different experimental situations. It is argued that the results of such experiments give reliable indications of ITD values at which acoustic-image trajectories cross the centre axis. It will be seen that experiments which were conducted using both image lateralization and image centering methods gave results which showed closely similar ITD's for judged centered images.

Listeners were required to adjust the ITD parameter to bring to centre position the various acoustic images which could be identified. The control was sufficient to vary ITD over a range of 4 msec (more than adequate for most experiments). Listeners were encouraged to approach centre position from both directions an equal number of times.

The experiments were not elaborately instrumented. Listeners were seated where they could conveniently adjust the ITD control and yet have no reliable clue as to the actual ITD at any moment. The experimenter observed the value of ITD on the oscilloscope display in the usual manner, and recorded the magnitude when the listener reported a centered image. Signals were turned on and off at the listener's request.

7.3.0 Apparatus

Apparatus for the acoustic-image lateralization experiments was unchanged from the previous chapter. The various signal and masking manipulations were carried out as previously described (Sections 6.2.0, 6.3.0).

The model used for approximating basilar membrane displacement is shown in its simplified form in Fig. 7.1. Included in the figure are the component values used to simulate basilar membrane characteristics at points maximally responsive to various frequencies. The buffer amplifiers were transistor emitter-followers having unity gain. The amplitude factors were taken into account, where necessary, as multiplicative constants rather than as fixed gains in the buffer amplifiers. R, L and C components were laboratory decade units which could be easily and accurately set to the required values.

The acoustic transient waveform which appeared at the output of the artificial ear was passed through a high-pass filter and then applied to the input of the model. The purpose of the filter was to eliminate the very-low-frequency noise which arose in the output due to the effect of building vibrations, etc. on the sensitive condenser microphone in the artificial ear. The cutoff frequency used was normally 100 cps; however, when such filtering could cause degradation of the acoustic transient signal, the filter was either not employed or its cutoff frequency was reduced to a safe value. In the latter cases it was very difficult either to observe or to photograph the oscilloscope display because of the unstable trace.

7.4.0 Experiments

7.4.1 Further Experiments with Cophasic and Antiphasic Transients

These experiments were conducted using a variety of acoustic waveforms of an impulsive kind; judged trajectories of the various important images perceived were considered in comparison with the basilar membrane displacement patterns for corresponding transients and various membrane locations (as indicated by the membrane model). It will be seen that cophasic and antiphasic transients are of some value in experiments of this type.

The curious result was reported earlier, Figs. 5.15(c) and (d), which was not supported by a comparable experiment due to Flanagan (1962). Flanagan observed that with unmasked antiphasic transients, image centerings were achieved at about ITD = ± 400 µsec; however, the results of Fig. 5.15 indicate single image centre-crossings at closely ITD = 0 for both cophasic and antiphasic signal configurations.

With the greater background of experience available at this stage, it was decided to repeat this simple experiment using a different repetition period (i.e., 20 msec) of a value such that the tonal harmonic images were less obtrusive. Fig. 7.2 illustrates the result of this interlaced experiment, using both cophasic and antiphasic transients. The listener, after some effort was able to identify three impulsive images: one of high-pitched character, one of low-pitched character and one of a character which was reported as "perhaps extremely low pitched." The latter image was extremely difficult to find, and the listener used in this experiment was the only one apparently able to effect the resolution. The other three listeners were tested briefly and gave results in the ITD = 0 region which, in all cases and for both high-and low-pitched images, resembled closely the judged trajectories for the high-pitched impulsive image illustrated here.

It will be noted in this result that judged trajectories for the high-pitched impulsive image were not seriously altered by the polarity modification. This effect was also seen in Fig. 5.15. The judged trajectory for the low-pitched impulsive image exhibited some rather interesting characteristics in the antiphasic condition, viz., a peculiar irregularity in the ITD = 0 region with consistent reporting of image dispersion. The latter observation applies also to the judged trajectory of the low-pitched impulsive image shown in Fig. 5.15 (antiphasic condition). It could be inferred from these observations, and from the general shape of the judgement curve, that two low-pitched images (having trajectories symmetrically disposed about ITD = 0) were simultaneously present. Positional judgements in this region could, therefore, relate to one or the other of these images, or to some intermediate position indicating, perhaps, a balance of attention between the two of them. If so, the low-pitched image identified in this experiment could well be related to the image doubly centered in Flanagan's experiments.

To clarify this point, experiments were arranged using signals found previously (Fig. 6.20) to give rise to a subsidiary image trajectory widely displaced from that of the main image, and clearly in evidence. The result, in Fig. 7.3, was achieved using binaural single repetitive transients bandpass filtered to 400-800 cps, which were employed in the four possible polarity modes. Measured acoustic waveforms are shown in Fig. 6.17. Both cophasic and both antiphasic results were very similar, with the antiphasic curves each displaying the irregularities and dispersed-image reportings as previously seen. The cophasic results may be seen to resemble those shown earlier in Fig. 6.21(c), and also in Fig. 6.20(a) where the S_1 and S_2 image trajectories were well separated. Only a single impulsive image was identified.

Other experiments along similar lines were conducted using transients with different bandwidths in each channel, in order to extend information on the subsidiary images, these being more clearly in evidence in experiments using antiphasic transients.

Results of typical experiments of this type are illustrated, for two listeners, in Figs. 7.4, 7.5, 7.6 and 7.7. These experiments used unfiltered transients in the left channel and transients bandpass filtered 20-400 cps in the right channel. Again, the relative signal polarity appeared to be the only important factor, with cophasic and antiphasic results different but comparable for each listener. Judgement curves for the two listeners were of basically similar form, but differed substantially in the spacings of image trajectories in the antiphasic experiments; F.E.T. indicated spacings of 1.2 and 1.35 msec, and D.G.P. indicated spacings of 1.95 and 2.1 msec. These results will later be discussed in detail.

Next, a number of exploratory experiments were conducted using signals which were selectively noise masked and filtered. Here, listeners were simply required to bring to the subjective central position, by adjustment of the ITD, all images which could be identified.

Initially, the wideband signals were symmetrically noise masked with high-pass noise having various cutoff frequencies. Image centerings were reproduced with a high degree of accuracy, and in the antiphasic condition were closely grouped at values approximately symmetrical about ITD = 0. One listener identified and independently centered two impulsive images in certain instances. These were recorded independently. Results of these experiments are displayed in Table 7.1.

With binaural high-pass filtered masking noise the image centerings tended to be spaced by amounts which could be related to the cutoff frequency of the noise. One listener (F.E.T.) apparently continued to lateralize the low-pitched image which was identified in the 600 cps cutoff frequency condition; it was, however, reported as the less important image. The other listeners could not confidently resolve this image in the higher cutoff frequency conditions; however, one of them (D.G.P.) reported a few isolated centerings at ITD values which were appropriate to the lower-pitched image. Because these centerings were few in number, and were reported with some uncertainty they were not included in the results.

The results with low-pass noise confirmed earlier observations that the high-pitched impulsive image was substantially independent of pulse polarity in respect of ITD for a centered image.

The above findings were instructive in two respects. Firstly, listeners could with relative ease, bring to the judged centre position, images which in the usual image-lateralization experiments were not judged as crossing the central axis. Secondly, the sustained centering of the lower-pitched image by the listener F.E.T. seemed to parallel the earlier (Fig. 7.2) lateralization of a third impulsive image. The former phenomenon had also been observed earlier and, in fact, prompted the training procedures adopted in certain experiments of the previous chapter (Section 6.4.1).

As a test of these findings two further experiments of the image-lateralization type were conducted. The first used wideband antiphasic transients with and without 600-15000 cps bandpass filtered masking noise in both channels. The listener attempted to lateralize the same image throughout these interlaced experiments. Fig. 7.8 illustrates the result, which shows very clearly that this image is undeniably present under wideband, unmasked stimulation, albeit very difficult to identify in that circumstance. It is to be noted that the ITD values at which the judged image trajectories crossed the centre axis were similar for masked and unmasked conditions, and further, that these correspond closely with the image-centering values of ITD recorded in Table 7.1.

The last image lateralization experiment also used the signals of an earlier image-centering experiment. Binaural masking noise bandpass filtered 1000-15000 cps was bilaterally applied along with antiphasic, unfiltered transients. Fig. 7.9 illustrates the result of this experiment in which the listener attempted to lateralize the same two images that he reported in the corresponding centering experiment (Table 7.1). It is clear from the form of the judged trajectory, that the lower-pitched image is that which was lateralized in the experiment of Fig. 7.8 and centered in the corresponding image-centering experiment (Table 7.1). The higher-pitched image trajectory is less well defined because no center-position judgements were reported. However, it is believed that the judgement curve represents portions of two symmetrically disposed image trajectories, and on this basis it may be estimated that, should these trajectories parallel those of the lower-pitched image, they would cross the centre axis at ITD's of about ± 0.5 msec. The higher-pitched image of the corresponding image-centering experiment (Table 7.1) was centered by this listener at ITD's of -0.55 msec and +0.5 msec.

Piecing the evidence together, it would seem that in experiments using wideband acoustic transients, it is possible to identify three impulsive images having different lateralization characteristics. In respect of ITD values for centered images, the impulsive image of highest pitch would seem to be largely independent of pulse polarity. Certain results could have led to a similar interpretation for the easily identified low-pitched image, see Fig. 7.2(b) and Fig. 5.15, but later findings (Table 7.1 and Fig. 7.9) suggested that positional judgements in these earlier experiments may have indicated a balance of attention between two images following trajectories close to and symmetrically disposed about ITD = 0.

With the intention of further elucidating the processes underlying the multiple impulsive images, a number of further image-centering experiments were conducted (Tables 7.2, 7.3, 7.4). However, because of the specialized nature of these experiments the description will be included in the discussion in Section 7.5.3.

7.4.2 Measurements on the Model for Approximating Basilar Membrane Displacement

The results just reported were next considered in the light of basilar membrane responses to the respective acoustic signals as indicated by the model previously described. The following description will cover only general features of the measurements; specific details and the comparative analysis will be considered in some detail in the discussion.

The initial measurement used wideband (20-15000 cps) acoustic transients from the artificial ear as an input signal. The output of the model was recorded when component values were set to simulate basilar membrane characteristics at various frequency (distance) coordinates - Fig. 7.10 shows photographed traces representing the acoustic transient (the input to the model) and basilar membrane displacement waveforms (the output of the model) at points maximally responsive to 200 cps, 500 cps, 1000 cps and 1500 cps. The scale factors are shown for each of the traces , but it is to be noted that the amplitudes as shown are uncorrected. When the required amplitude correction factor (Fig. 7.1) has been applied, (a rise of 5 dB/octave below 1000 cps - flat above this frequency) the amplitudes of the first three peaks of the recorded displacement waveforms were seen to vary as shown in Fig. 7.11.

The apparent low-pass filter action of the cochlea is clear from the progressive slowing down of the displacement oscillations as locations further towards the apex are considered.

The next step was to observe in similar fashion the membrane responses to a filtered acoustic transient, in this case, a signal bandpass filtered 20-600 cps. Figs. 7.12 and 7.13 show~~s~~ the waveforms of the filtered acoustic transient, and of the basilar membrane displacement at the 1500 cps, 1000 cps, 700 cps, 600 cps, 500 cps and 200 cps points. Fig. 7.14 illustrates corrected displacement peak amplitudes at 100 cps intervals over the same frequency range.

It is apparent that the displacement waveform retains the temporal characteristics of the acoustic transient in a relatively unmodified form down to the point in the apical region which is selective to the highest spectral component of the signal. Thereafter the expected apparent low-pass filter action is seen to occur.

The corrected amplitude curves are enlightening as well. It can be seen that the maximum basilar membrane displacements occurred in the 600-1000 cps region with the 20-600 cps bandpass filtered signal whereas with the unfiltered transient these maxima were seen to occur in the 1000-1500 cps region.

These findings are in general agreement with the assumptions regarding basilar membrane displacements in response to acoustic transients which were stated in Section 6.5.0.

It is pertinent also to consider a few situations which relate to image-lateralization experiments of special interest. Basilar membrane responses for the widely used repetitive pulse versus pulse pair signal configuration are of obvious importance. Fig. 7.15 illustrates the acoustic waveforms for two unfiltered transients separated by 1.4 msec (δt = 1.4 msec, T = 20 msec). There would appear to be little interaction between these transients. Other illustrations show waveforms of basilar membrane displacement at the 500 cps point in response to acoustic transients separated by different δt intervals. At δt = 2.8 msec the membrane responses to the two transients appear to be substantially independent. At δt = 1.4 msec the two responses interact in such a way as to produce a smooth oscillatory waveform with a substantially enhanced condensation peak. At δt = 0.8 msec the interaction is more complex, with displacements only poorly reflecting the presence of the two transients. The last illustration shows the δt = 0.8 msec information clearly in evidence in the displacement waveform at the 1000 cps point on the membrane model.

It would seem that, within limits, there is a region of the basilar membrane which responds in a smoothly oscillatory manner, and maximally for each magnitude of pulse-pair separation. The curves of Fig. 7.16 further illustrate this point. Here are plotted for δt = various, the peak amplitudes of basilar membrane displacements at the 500 cps point. Clearly, there was only one maximum at which the δt information is reproduced or represented in the displacement waveform (the other occurred when the two transients were superimposed). The important maximum related to δt values in the vicinity of 1.6 msec. See also Fig. 7.15.

Also bearing on this matter are the curves of Figs. 7.17 and 7.18 which illustrate, for $\delta t = 2.0$ msec and $\delta t = 3.0$ msec, corrected amplitude responses at points along the basilar membrane; the signal was band-limited to 20-600 cps. It may be seen, in the curves for the $\delta t = 2.0$ msec situation, that responses in the 400-600 cps region were increased in amplitude compared with those for the single pulse (Fig. 7.14). In the $\delta t = 3.0$ msec situation responses in the 300-500 cps region were increased in amplitude, again relative to the curves of Fig. 7.14.

Results of a not dissimilar type are shown in Fig. 7.19. Here, the displacement waveforms at the 300 cps point are shown for unfiltered repetitive single transients applied at different repetition periods. Waveforms are illustrated for T = 8.0 msec, T = 3.0 msec and T = 1.5 msec. It is evident that at the longest repetition period the responses to individual transients were completely independent. At T = 3.0 msec successive responses appeared to be largely additive, resulting in an almost uniform oscillation. At the still smaller repetition period, the displacements were still periodic at the signal rate, but with diminished amplitude.

The experimental results illustrated in Figs. 7.4 to 7.7 were selected for special consideration and the following relevant waveforms were obtained using the model. Fig. 7.20 illustrates waveforms for the unfiltered and 20-400 cps bandpass filtered acoustic transients and for the displacement waveform of the basilar membrane at the 400 cps point in response to each of these signals.

7.5.0 Discussion

7.5.1 A Note on Cochlear Neurophysiology

The study of basilar membrane displacements alone is insufficient to enable unambiguous interpretation of experimental results. It is first necessary to have some knowledge of, or make some assumptions about, the manner of mechanical to neural transduction. It is widely accepted that neural firings arising from the organ of Corti as observed in the cochlear branch of the auditory nerve, do so as a direct consequence of basilar membrane displacement. These displacements are held to set up a shearing motion between the reticular lamina and the tectorial membrane, thus bending the tiny hairs which are suspended between these bodies. Associated with each tuft of hairs is a hair cell which, according to one interpretation, may initiate neural action potentials as a consequence of the bending. The actual mechanism which effects the neural discharge is not yet understood, but it is believed to be related to the cochlear microphonic voltage (Davis, 1957, 1958, 1961). In any event, there would seem to be substantial support for the view that neural action potentials associated with the outer hair cells (the more sensitive cells) are initiated at some point on the unipolar upward (in the direction of the tectorial membrane) deflections of the basilar membrane. This direction of displacement would be caused by an outward movement of the stapes such as would arise from an impulse of pressure rarefaction. Flanagan (1962) first used this assumption to interpret the results of psychophysical experiments. Since then, Harris, et al. (1963) and Nordmark (1963), among others, have continued to support this view. The assumption, which was based on the findings of von Békésy (1953) and Davis (1958), has since been given further physiological support by the findings of Kiang, et al. (reported in Weiss, 1964).

Kiang, et al. observed single fiber responses to pure tones and clicks presented through condenser earphones and, inter alia, considered responses to clicks of different polarity. These workers found that PST (post stimulus time) response histograms for single fibers of the auditory nerve showed, comparing response to condensation with those to rarefaction clicks, differences in the times of occurrence of peaks corresponding to $1/(2f_0)$; where f_0 is the frequency of lowest threshold of firing for a given fiber. The first peak in the PST histogram in response to a rarefaction click was found to lead the first peak in the PST histogram for the response to a condensation click. Interpretation of this finding requires a brief consideration of basilar membrane responses to acoustic transient stimulation.

Although the waveform of the actual acoustic transients used by Kiang, et al. is not known, a condenser microphone was used and it is assumed that they were basically of the form of the transient shown in Fig. 7.10 (only a substantial difference would affect the argument which follows). It may be seen in the waveforms of Fig. 7.10 that a given point on the membrane (according to the model) exhibits multiple displacement peaks at spacings which may be associated with the period of the sinusoidal acoustic signal to which that point is maximally responsive. On the reasonable assumption that the nerve fibers associated with a location on the membrane exhibit minimum firing thresholds at the frequency which produces maximum mechanical response at that point, the findings of Kiang, et al. may be interpreted.

For example, consider the basilar membrane displacement waveform at the 1000 cps point (Fig. 7.10). For the illustrated (condensation) signal configuration an auditory nerve fiber associated with the 1000 cps point on the membrane would be expected to show activity only on the second and fourth peaks of membrane displacements (i.e., those which are displaced towards the tectorial membrane, as for a pressure rarefaction). However, if the acoustic signal is inverted (i.e., becomes a rarefaction transient), the membrane displacements are also inverted, and in the same nerve fiber, firings would be anticipated for the first and third (rarefaction) peaks of displacement. It will be noted that the latter (rarefaction transient) firings would lead the former (condensation transient) firings by about 0.5 msec, i.e. $1/(2f_0)$, where $f_0 = 1000$ cps. The same proportional difference would be expected to hold for other locations (and other nerve fibers) on the basilar membrane.

In short it seems that the assumption that neural firings occur only on displacements of the basilar membrane which may be associated with pressure rarefactions, is entirely reasonable. This principle will be employed in the discussions which follow.

It is not presumed that this is the only cochlear mechanism for initiating neural action potentials. There would appear to be at least one other source - the inner hair cells - which are apparently sensitive to the longitudinal gradient of basilar membrane displacement (Flanagan, 1962, and references cited therein).

For the purposes of the present discussion, however, only the activity of the more sensitive outer hair cells will be considered.

On the basis of the view of mechanical-to-neural transductions outlined above, it is clear that neural activity may now be related to temporal features of basilar membrane displacements, to position of origin along the cochlea, and ultimately to the acoustic waveshape.

7.5.2 Current Views on Binaural Fusion and The Associated Cross-Comparison

It is useful to recapitulate the broad outlines of the currently accepted working picture of the mechanism in which the binaural signals are believed to interact. The foundation of such a view is the concept of binaural fusion, i.e., combination of aural signals in the two (left and right ears) sensory channels to give rise to the perception of sound images. The actual form of the cross-comparison is not yet known; however, it is possible to lay down a few basic requirements for the mechanism of cross-comparison and the processes closely associated with it.

It is clear that in the mechanism of cross-comparison there must be some form of equivalent interaural time-to-space transformation. This is required in order to represent the listener's experience of auditory space as objectively assessed and to the extent relevant here, i.e., that perceived lateral displacements of binaurally created images can be associated with ITD. Various descriptive models have been proposed by other workers and will now be considered briefly.

Sayers and Cherry (1957) proposed a time-to-space transformation as effected by the expedient of a real time multiplicative combination and integration (i.e. a running crosscorrelation) of the signals from the two ears, followed by the assumption that interaural amplitude and time differences are spatially represented on a conceptual surface. In such a representation ITD appears as a displacement in (correlation delay), and hence in conceptual space. Questions of detail in the transformation, and thereby many important characteristics of perception, were entirely avoided.

In Sayers (1964), matters pertaining to the extent of image displacement were introduced and handled similarly except that the limitations of dichotomous (simple 'left' or 'right' of centre) judgements were no longer involved. Using this hypothesis, it is easy to obtain a linear relationship between predicted image displacement and ITD (although no special significance need be attached here to this observation). It has already been shown that such a linear relationship between ITD and judged image position can be readily achieved in experiments; for experiments of this and the present type it is convenient to use a technique such that a linear ITD (time)-to-position (space) transformation is operative (see discussion of this point in Section 3.2.1.

A related scheme was proposed earlier by Jeffress (1948), and has been incorporated into the "triplex" theory of pitch perception as described by Licklider (1959). The theory assumes that there is a mechanical frequency analysis in the cochlea and that the products of the analysis are subjected to a twofold coincidence or correlation analysis. One of these analyses utilizes the ordered neural outputs from the cochlea (mapped according to frequency) in a mechanism which transforms ITD directly into spatial position. The basic elements of the mechanism of cross-comparison are neuronal 'delay-lines' which have their origins in the two cochleae, and pass each other in parallel fashion. Between adjacent lines (from corresponding points on the two cochleae) are neural elements which are sensitive to temporally coinciding excitations at the points on the lines to which they are attached. The effect of the coincidence mechanism is to focus activity within a vertical plane between the cochleae, which has a position dictated by the interaural time relations between the signals at the two ears. Identical binaural signals would produce activity (i.e. coincidence) at the centre position. Binaural signals with ITD would produce activity at a point displaced from centre; the amount and direction of displacement would be directly related to the magnitude and polarity of the ITD.

This already schematized model may be further simplified for the purposes of this discussion, in the manner of Fig. 7.21(a). Here are illustrated two of many such proposed delay-line elements, one from each cochlea. It is hypothesized that volleys of neural activity associated with some significant feature of the binaural acoustic signals (e.g., the rarefaction peaks of a sinusoidal pressure waveform) are launched along these identical delay-lines and pass at points determined by the repetition period of the binaural signal and the ITD, thus effecting the required interaural time-to-position conversion. In the sequence of illustrations it may be seen that as ITD is increased from zero to T (the repetition period of the sinusoid) a pattern of neural coincidences is established. Within the region designated as the 'conceptual space' (i.e., a factor which parallels the perceived auditory space in which binaurally created images are established and move), the point of neural coincidence may be seen to move from the centre position as ITD is increased from zero, reaching the lateral extreme position at ITD = T/2, at which stage it is joined by and subsequently replaced by a similar neural coincidence at the contralateral extreme position which moves to the centre position for ITD = T = 0. Naturally, such a pattern would be repetitive in ITD for a repetitive signal such as a sinusoid.

If in the above description one reads "image position" instead of "point of neural coincidence" the parallel between the descriptive model and the image lateralization results for binaural single component tones (Sayers, 1964) is clear. In short, on this view, lateral position of neural coincidence is associated with the perception of a binaural sound image in a corresponding position.

The choice of the extent of the 'conceptual space' in the diagram was convenient for purposes of introduction; however, it is clear that if the transmission rates of the delay-lines are constant (and equal), binaural signals of higher repetition frequencies would give rise to similar patterns of image movement with ITD, but in which the lateral maxima of image displacement would diminish with increasing frequency. Sayers (1964) considered this point and described confirmatory experimental evidence.

If, on the other hand, the signal frequency were to be decreased the image would be displaced to the limit of the conceptual space at ITD's progressively less than T/2, whereupon this region of the model no longer accommodates an image. Experimental results suggest, however, that in such an event the image is judged to remain at the lateral extreme position and eventually to split into symmetrically disposed elements; see for example results of experiments employing binaural single repetitive transients (Fig. 4.3) and the discussion in Section 4.5.0.

A descriptive mechanism of this form may also be seen to accommodate results of experiments employing single-versus double-repetitive transients and the S_1 and S_2 impulsive images which were hypothesized to arise.

It is not intended to pursue the analogy further, but rather at this stage to consider a few of the more general aspects of the mechanism. Arising out of the present work are further features which must be incorporated into the working picture. Firstly, incoming binaural acoustic signals are apparently subjected to selective processing whereby some signals are cross-compared so as to produce fused images, while others apparently are not (see point 1.0 of the summary in Section 6.5.0). Secondly, there must be provision for the establishment of multiple-and split-images (see Chapter 5).

The descriptive mechanism as developed so far, stands in a very elementary form. The consideration of acoustic signal waveform alone, was a concession to simplicity, and useful for purposes of introduction. A realistic mechanism, however, must consider actual neural firings in response to basilar membrane displacements. Clearly the characteristics of basilar membrane displacements will reflect not only properties of the acoustic signal but also the dynamic properties of the cochlear partition. This would be especially important in the case of wide-band acoustic signals, e.g., unfiltered transients. It is therefore of fundamental importance to establish the relation between the neural activity which is important in the cross-comparison, and the waveform of the acoustic signal.

Leaving aside the question of the actual site of the operation there would seem to be four logical modes for the cross-comparison process, viz.,

(1) the interaction of neural impulses that are closely synchronized to certain temporal features of the acoustic waveform (such activity might arise from the basal turn of the cochlea),

(2) the interaction of combined neural impulses arising from the entire length of the cochleae (i.e., presumably the action potentials of the whole auditory nerves, see Section 6.2.0),

(3) the interaction of neural information arising from the maximally responsive or other specially defined regions of both cochleae,

(4) the interaction of neural impulses arising from corresponding portions of both cochleae.

The question to be answered is: which, if any, of these possibilities provides the closest parallel to the actual mechanism of cross-comparison?

There are a few important experimental observations which are of direct relevance in the resolution of this problem. For example, in Chapter 5 it was shown that with binaural repetitive transients, tonal and impulsive images could be simultaneously identified and lateralized. The tonal images clearly arose from a spectral selection or analysis of the binaural repetitive transient signals; the impulsive images were apparently related to some dominant features of the acoustic waveform. These findings argue against the first suggestion in respect of the tonal images, although not in respect of the impulsive images.

However, in later experiments it has been shown that multiple impulsive images may be simultaneously perceived, and that these appear to be associated with neural activity in two regions of the cochleae; the high-pitched impulsive image with basal region activity, and the low-pitched image with apical region activity. This finding alone is sufficient to cast some doubt on the indication mentioned in the previous paragraph. Near the end of Section 6.5.0 there was a discussion of further evidence which argued against the consideration of acoustic waveform features alone. However, further clarification was required, and a discussion of suitable experiments reported earlier (Section 7.4.0) is now appropriate.

7.5.3 Discussion of Experiments Employing Cophasic and Antiphasic Transients

A definitive experiment to clarify these issues would be one using different acoustic transient waveforms in each channel, which produced curves of impulsive image lateralization judgements versus ITD that would seem to allow interpretation in only one way. The results shown in Figs. 7.4, 7.5, 7.6 and 7.7 illustrate such an experiment. Here, for two listeners, are shown judgement curves resulting from experiments in which the left-ear signal was unfiltered and the right-ear signal was filtered to a pass-band of 20-400 cps. Figs. 7.20(a) and (c) illustrate waveforms of these acoustic transients as measured using an artificial ear. These signals were used in the two cophasic and the two antiphasic configurations. Judged trajectories of the single low-pitched image which was identified, show indications of a number of image trajectories in both cophasic and antiphasic configurations, although these were most clearly evidenced in the antiphasic conditions. Spacings of these trajectories in ITD were consistently about 1.2-1.35 msec for one subject (Figs. 7.4 and 7.5) and about 1.95-2.1 msec for the other (Figs. 7.6 and 7.7).

Using the measured acoustic waveforms of Figs. 7.20(a) and (c) it is possible to construct predictive curves of image movement as a function of ITD on the basis which follows. Considering temporal information relating to rarefaction peaks only (as justified in 7.5.1) it is found that the judged trajectories, which were closely similar in both cophasic and both antiphasic conditions, cannot be immediately accommodated. This can be simply tested by deducing from the waveforms the ITD which would be required to bring into alignment the rarefaction peaks of the contralateral acoustic transient waveforms (that of Fig. 7.20(a) in the left ear, and that of Fig. 7.20(c) in the right ear) when employed in the various polarity configurations. This procedure results in ITD values at which centered images would be anticipated.

As an example of the disparity between judged image centerings and those predicted on a complete waveform basis, consider the two cophasic signal configurations. If both signals are applied so as to produce initial peaks of pressure condensation (as in Figs. 7.20 a and c), the ITD value at which a centered image would be anticipated is seen to be 1.6 msec (i.e., the interaural time between peaks B and H), left side (unfiltered transient) lagging. If both signals are inverted, peaks A and G would be the dominant factors, and the ITD for an anticipated image centering is about 0.8 msec (left lagging). The disparity between these predictive values (1.6 msec and 0.8 msec) is rather large when compared with the close similarity of the actual results obtained with both listeners. These were 1.1 and 1.05 msec for one listener (Figs. 7.4, 7.5) and 1.25 and 1.15 msec for the other listener (Figs. 7.6, 7.7).

If, however, one considers basilar membrane displacement waveforms at the 400 cps points of both cochleae, a rather different picture emerges. Figs. 7.20 (b) and (d) illustrate for each of the acoustic transients, the approximate basilar membrane displacement waveforms at the point maximally responsive to 400 cps. It will be seen that, in spite of the radically different acoustic transients in the two channels, the waveforms of membrane displacement at the 400 cps points have much in common, viz., the spacing of successive unipolar displacement peaks: in the wideband case about 2.2 msec - 2.6 msec and in the filtered case about 2.6 msec - 3.0 msec.

It would be entirely inappropriate to claim a very precise quantitative relation between actual neural responses to these acoustic transients and the assumed neural concomitants of membrane displacements as indicated by the basilar membrane model. However, bearing this in mind, an analysis on the basis of the approximate membrane displacement waveforms is instructive.

On the basis that cross-comparisons are effected between neural impulses arising from the 400 cps regions of both cochleae it is possible to estimate from the relevant waveforms of basilar membrane displacement the ITD values at which centered images would be anticipated. With this information it is possible to construct predictive image trajectories and to compare these with judged image trajectories. In Table 7.2 are recorded ITD values for predictive image centerings as estimated from waveforms of basilar membrane displacement at the points maximally responsive to 400 cps (Figs. 7.20 b and d). Letter designations accompanying the predictive values refer to the relevant contralateral rarefaction peaks of membrane displacement. Fig. 7.22 illustrates predictive image trajectories constructed on the basis of these estimations. On the supposition that images arising from cross-comparisons of neural impulses associated with initial or large rarefaction peaks of basilar membrane displacement may be of special importance, these images have been denoted by underlining of the relevant ITD values, or by special emphasis of relevant plotted trajectories.

It is clear from a comparison of the judged trajectories in Figs. 7.4-7.7 and the predictive trajectories of Fig. 7.22, that the lateralization characteristics of the low-pitched image perceived in these experiments may be accommodated reasonably well in terms of a cross-comparison of neural impulses arising from the 400 cps regions of both cochleae. It will be seen from comparison of the 'measured' and 'predicted' ITD values for centered images (Table 7.2) that the picture is equally as clear. In subsequent comparisons of this type, predictive image trajectories will not be sketched and comparisons will be made between image-centering ITD values.

In this case, the decision to study responses at the 400 cps point on the basilar membrane model was based broadly on the observation that the greatest displacements of the membrane to a 400 cps low-pass filtered transient occurred near the 400 cps point (see Fig. 7.14, which shows a peak of displacements in a region related to the low-pass cutoff frequency). It might be argued that the dominant neural contributions to the relevant cross-comparisons arise in the vicinity of this point of maximum response (see Flanagan 1962, and Harris, et al., 1963).

The latter point raises the question of a possible explanation in terms of a cross-comparison of neural impulses arising from the maximally responding regions of both cochleae (the third suggested mode of cross-comparison). On the basis of Fig. 7.11 it would seem that the important neural information might arise from **the** 1500 cps region of the cochlea receiving the unfiltered transient (the left-ear channel). Using the waveform of basilar membrane displacement at this point (Fig. 7.10) and the waveform of basilar membrane displacement at the maximally responding (400 cps) point of the contralateral membrane (Fig. 7.20d) it is again possible to compute predictive ITD values for centered images. However, in this case a new factor must now be considered. As was stated earlier (Section 7.2.0) the delay-line portion of the basilar membrane model was omitted. Where cochlear responses at non-corresponding points are of importance, the delay factors must be incorporated into the computations. The waveforms illustrated in Fig. 7.23 are tracings of those shown earlier. Here are shown symbols designating the waveform peaks and the times at which the various peaks occur. These times are measured from the onset of the electrical signal pulses (between which the experimental ITD's were measured) and include the computed delay factors. Thus it is possible to ascertain interaural time relations between peaks of any two of the illustrated waveforms by algebraic addition of the appropriate numerical values.

Based on displacement waveforms at the 1500 cps point of the left-ear membrane (Fig. 7.23a) and at the 400 cps point of the right-ear membrane (Fig. 7.23f) the predictive values for centered images were computed and are tabulated in Table 7.2. Comparison of these values with the judged image centerings included in the same table reveal no consistent pattern of similarity, as existed in the case of the previous predictive values. Not only was there no real agreement of the values in any one of the experimental conditions, but the predictive values in the two cophasic and the two antiphasic situations were quite dissimilar; however, judged image trajectories were closely similar in the corresponding conditions. It appeared, therefore, that the important cross-comparison in this case was not effected on neural impulses arising from maximally responding regions of both cochleae.

The remaining mode of cross-comparison, which was suggested earlier, would also appear not to be adequate to describe the results of these experiments. The clarity with which the two image trajectories were reproduced in judgement patterns implies the cross-comparison of neural firings from each cochlea which are closely synchronized and at similar intervals. The combination of neural impulses arising throughout the length of the cochlea would be expected to produce a pattern of neural impulses which reflects to varying extents a wide range of intervals. Waveforms shown in Fig. 7.10 exhibit a portion of this range of intervals between successive unipolar peaks (and hence, bursts of neural activity). If the combined neural activity was dominated by the maximally responding regions of the basilar membrane, the picture changes to that of the third suggested mode of cross-comparison which has been discussed and provisionally discarded on the basis of evidence presented in the preceding paragraph. If the various portions of the cochlea contribute to similar extents to the combined neural activity, the resultant pattern would be far from closely synchronized.

It is concluded, therefore, that the only satisfactory description of the experimental results of Figs. 7.4 - 7.7 is provided in terms of a cross-comparison of neural activity arising from corresponding regions of both cochleae, in this case the regions maximally responsive to about 400 cps. The working picture of the mechanism of cross-comparison may now be extended to incorporate this feature. Fig. 7.21(b) illustrates the elements of cross-comparison as discussed above and illustrated in Fig. 7.21(a), coupled in parallel fashion between the two uncoiled cochleae. When the neural impulses have passed through the cross-comparison elements they are shown as being passed on to higher centres for further processing. Each pair of delay-lines (with associated coincidence-detection elements) is hypothesized as being capable of contributing to the perception of a fused sound image; thus it is possible for a number of images to coexist, each reflecting properties associated with the regions of the cochlea from which the relevant neural impulses arise. The overall pattern of images, as viewed within the conceptual space, might therefore be quite complex for binaural wideband signals such as acoustic transients. In such a case the lateralization properties of images associated with neural impulses arising from different regions of the cochlea may be expected to reflect to an important extent the dynamic properties of the basilar membrane in the relevant regions. It is hypothesized that the images which dominate the listener's impression are in a simple way associated with regions of the basilar membrane which are either maximally displaced or show peaks of response relative to neighbouring regions.

It is now interesting to further examine the experimental results of this and the previous chapter with a view to establishing the validity and utility of this tentative working picture of the mechanism of cross-comparison.

In direct support of the image lateralization results just discussed (Figs. 7.4-7.7) are the results of image centering experiments which are recorded in Table 7.3. These experiments used the same signals as were used in the earlier experiments, but with the channels reversed so that the 20-400 cps filtered signal was in the left-ear channel, and the unfiltered signal was in the right-ear channel.

As well as the four polarity configurations previously used, this set of experiments also included a sequence in which the right-ear (unfiltered) signal was selectively masked. The aim of this sequence of experiments was to further test the hypothesis of cross-comparison between corresponding cochlear regions. If, as was proposed earlier, the cross-comparison is effected between neural impulses arising only from corresponding narrow regions of the cochleae, lateralization properties of the impulsive image would not be modified by masking of neural activity arising, in either cochlea, outside of the contributory regions. If, on the other hand, the cross-comparison is effected between those points of the membrane which respond with greatest amplitudes (and hence presumably generate the most important neural activity), the lateralization properties would be expected to alter as the points of cross-comparison are changed. As stated earlier the point of maximal response in the right-ear (unfiltered) cochlea would be near the 1500 cps point of the basilar membrane, and that of the left-ear (filtered) cochlea would be near the 400 cps point (see Figs. 7.11 and 7.14). However, high-pass filtered masking noise added to the right-ear (unfiltered) signal would be expected to largely eliminate synchronous neural firings in the cochlear region extending from the oval window to a distal point which is maximally responsive to the cutoff frequency of the filter. If this cutoff frequency is less than about 1500 cps, it would be expected that the portion of the cochlea giving rise to the most important neural information would be that maximally responsive to frequencies just beyond the lower cutoff frequency of the noise (again, attention is directed to Fig. 7.11, and particularly to the monotonic decline of displacement amplitude from near the 1500 cps point to the apex). On this basis, the masking of the right-ear signal would force the cross-comparison to be effected between the 400 cps point on the left-ear cochlea and points on the right-ear cochlea associated with the cutoff frequency of the high-pass filtered noise. In the present experiments cutoff frequencies of 1000 cps, 600 cps, 500 cps and 400 cps were used, which, on the basis of the displacement waveforms shown in Fig. 7.23(b), (c), (d) and (e) should produce substantially different image lateralization characteristics in cross-comparisons with neural concomitants of the displacement waveform of Fig. 7.23(f).

Included in Table 7.3 are results of the image centering experiments and, for each, predictive ITD values estimated using the waveforms of Fig. 7.23, for the two important suggested modes of cross-comparison: that based on a cross-comparison of neural activity arising from corresponding regions of the cochleae, and that based on cross-comparison of neural activity arising from maximally responding regions of the cochleae.

It will be seen that these image-centering results are closely similar to the corresponding image lateralization results of Figs. 7.4 - 7.7 which were recorded in Table 7.2 (polarities of ITD are reversed because the signals were interchanged). In all cases the results may be seen to be best described in terms of a cross-comparison of neural activity arising from corresponding regions of the cochlea. The agreement extended also to the masking experiments where image centerings were reported and predicted at similar ITD's for all masked conditions. This observation adds further support to the proposed working picture of the mechanism of cross-comparison.

A confirmatory experiment of similar type was arranged using unfiltered transients in the right-ear channel and 20-600 cps filtered transients in the left-ear channel. In this case cross-comparison might be anticipated either between both 600 cps points, or between the 600 cps point on the left and the 1500 cps point on the right, according to which of the two suggested modes is considered.

In Table 7.4 (Experiment B) are listed, for two listeners, ITD's at which centered images were reported. Recorded also are predictive image centering ITD's estimated using the waveforms of Figs. 7.23 (a), (c) and (g) for the two methods of cross-comparison now under examination (the remaining two suggested modes of cross-comparison have been rejected on the same grounds as were discussed in the immediately preceding paragraphs). Once again the centerings predicted using the proposed working picture are arranged singly or in groups at ITD values which can be associated with those established in the psychophysical experiments. No real importance is attached to the observation that not all predicted centerings were detected by experiment. It would seem entirely reasonable that some images are of negligible proportions, or are spatially positioned such that they are considered in a 'compromise' image along with a similar-sounding image following a closely parallel trajectory. Such 'compromise' judgements have been discussed earlier (see Section 4.5.0 and 7.4.1) and have been described as indicating a balance of attention between two similar-sounding images.

The predictive centerings based on cross-comparison of neural information from maximally responding regions are seen to bear little resemblance to the experimentally determined values.

It is concluded, therefore, that for experiments using transient signals which produce maxima of activity at different points on contralateral cochleae, the results are best described in terms of a cross-comparison of neural information arising from corresponding points on both basilar membranes, (Tables 7.2, 7.3, 7.4). This was held to be the case also when the point of maximum neural response in one cochlea was presumably altered by the addition, to the appropriate signal, of high-pass filtered masking noise (Table 7.3).

The next results to be considered were achieved in experiments using repetitive single unfiltered acoustic transients in both channels. In these experiments a new problem was raised - that of accounting for the simultaneous presence of three images of impulsive character, having quite different lateralization characteristics.

In an early image lateralization experiment in the present chapter the three images were independently identified and lateralized in cophasic and antiphasic conditions (Fig. 7.2). It was quite clear in this result that judged trajectories of the high-pitched impulsive image were unmodified by the signal inversion in one channel; those of the commonly perceived low-pitched image, however, were altered but in a way which was not immediately clear. Judged trajectories of the third image (indicated by X's on the diagram) were also altered by the signal inversion but in an apparently well defined manner. The question of the lateralization properties exhibited by the low-pitched image was taken up in Section 7.4.1, and it was there concluded that the judged trajectory for this image as depicted in Fig. 7.2 (b) most likely represented a balance of attention between two images of similar character which were following trajectories close to, and symmetrically disposed about ITD = 0. In the resolution of this problem the results of image-centering experiments were found useful (Table 7.1).

These latter experiments were instructive also in a different respect. Bilateral masking of these unfiltered transient signals with symmetrically filtered noise having different high-pass cutoff frequencies enabled the number and lateralization properties of the impulsive images to be altered. This observation immediately suggested that these multiple impulsive images may be associated with neural activity arising from different regions of the basilar membranes. Where listeners did not discern these multiple images, it was clear from their reported image centerings that they had selected and tracked the dominant (higher-pitched) image reported by the one listener (F.E.T.) who had consistently reported multiple impulsive images (Table 7.1). It was further apparent that the spacing in ITD of image centerings reported for the dominant impulsive image in the antiphasic signal configurations were related to the cutoff frequency of the high-pass filtered noise; however, centering ITD's for the lower-pitched image did not alter appreciably from the condition where this image appeared in isolation (masked with 600 - 15000 cps filtered noise).

Before attempting to draw conclusions from these results, it is important to consider results of some further experiments which were basically similar to those just discussed but with the important difference that the masking noise was applied to one channel only. Results of these experiments are listed in Table 7.4 (Experiment C) and Table 7.5, for one cophasic condition and both antiphasic conditions. As in previous tabulations of this type, predictive image-centering values of ITD are included for comparison. Again, these predictive values were determined, using the waveforms of Fig. 7.23, for two of the suggested modes of cross-comparison: that on which the proposed working picture was based, and that in which the cross-comparison is effected between maximally responding regions. The remaining two suggested schemes were rejected because it was considered that neither a cross-comparison on a complete waveform basis (the transient waveform of Fig. 7.20a in both channels) nor that based on a combined cochlear nerve activity, could accommodate the clear-cut multiple-image centerings (Tables 7.1 and 7.5) which were reported. In fact, at this stage, it was almost certain that a single cross-comparison between maximally responding regions of the cochleae would not accommodate the results; however, this case was examined and has been included in the predictive results.

A number of interesting points emerge from an inspection of the tabulated experimental results of Table 7.4 (Experiment C) and Table 7.5. It is at once clear that the application of the filtered masking noise to one channel only produced image centerings closely similar to those reported earlier (Table 7.1) for experiments using identical filtered noise bilaterally applied (these results are reproduced in this table for comparison, and are shown as bracketed figures under those of the present experiments).

It is further apparent that, as in the results of Table 7.1, the listener who was not able to identify the multiple impulsive images, selected and tracked the single dominant image (the image of medium pitch reported by F.E.T.). There was one occasion (in the Antiphasic (2) condition) where this listener (B.McA.S.), on a second trial, produced image centerings clearly arranged in four groups which could be simply related to four of the five groupings reported by the second listener in the same experiment. The listener was, however, unable to assign pitch-like characters to these images which enabled the other listener (F.E.T.) to independently identify and centre (or lateralize) the various coexisting images.

With that background it is instructive to compare the results of image-centering experiments with those predicted on the bases discussed above. It is again clear that the ITD values predicted from cross-comparison of the assumed neural concomitants of basilar membrane displacements in maximally responding regions, will not accommodate the experimental results. The values do not exhibit any basic similarities in isolated conditions, nor do they reflect the similarities of the experimental results of the two antiphasic signal configurations. Moreover, the apparent identity between unilateral and bilateral masking conditions would not be predicted on this basis. For example, unilateral masking with 600-15000 cps noise would, on this view, force the cross-comparison of neural information from the 1500 cps (maximally responsive) point on the left-ear (unmasked) cochlea, and from the 600 cps (maximally responsive) point on the right-ear (masked) cochlea. With both listeners, results for both unilaterally masked antiphasic signal conditions were similar to each other and to the corresponding bilaterally masked case. There seems no reason to believe that these results could be accommodated by the ITD values predicted on this view (which were quite different in the two antiphasic configurations).

On the other hand, the ITD values predicted on the basis of the proposed working picture of the cross-comparison would appear to provide a more satisfactory description of the experimental results. A few of the salient points are discussed below.

The simultaneous perception of multiple impulsive images with different pitch-like characters would seem to be most satisfactorily accommodated by a descriptive framework which allows for multiple, substantially independent cross-comparisons from different portions of the cochleae. Clearly the proposed working picture of cross-comparison (Fig. 7.21b) is such a framework; however, it would seem necessary that there be a factor by which the importance of a certain number of cross-comparisons (i.e., possible images) are enhanced such that they influence the listener's experience sufficiently to be independently identified.

An apparently logical choice for one cross-comparison would be that which may be effected between neural information arising from the maximally responding region in one cochlea and that arising from the corresponding region of the other. This region would presumably be near the 1500 cps point (unmasked - maximally responding) or at a point associated with frequencies near the lower spectral limit of the masked bandwidth.

In the latter case the point of cross-comparison would be expected to alter with the cutoff frequency of the high-pass masking noise (for cutoff frequencies below about 1500 cps; see discussion of masking experiments reported in Table 7.3).

It will be seen that ITD's predicted on this basis agree reasonably well with those reported for the single dominant image (B.McA.S.) and for the image of 'moderate' pitch reported by F.E.T. The agreement is seen to extend also to the results listed in Table 7.1 (shown here) for the corresponding bilateral masking conditions. It is important to note that the judged similarity of both antiphasic signal configurations is broadly indicated by the predictive ITD's.

Extrapolating from these results, it was proposed that the lower-pitched image reported by F.E.T. may have arisen from a cross-comparison of neural information arising from more distal points on the basilar membrane which were in all cases apparently unaffected by the masking signals. Predictions based on displacement waveforms at the 400 cps points of both basilar membranes were selected as providing reasonable fits with experimental results for this image, and have been tabulated here. Again, predictive ITD's for both antiphasic signal conditions are seen to be appropriately similar.

No predictive values are shown which relate to the high-pitched impulsive image which was identified only in the unmasked conditions. It was assumed that this image is associated with neural activity arising from the basal regions of the cochlea. This assumption was based on earlier findings (Section 6.4.5) and present evidence which indicated that bilateral selective masking of neural activity in the basal regions prevented identification of the high-pitched image.

A curious feature of the high-pitched impulsive image as seen in the present results and in earlier results (Figs. 5.15, 7.2 and Table 7.1 - 20-1500 cps masked condition) was that it was judged to behave similarly in both cophasic and antiphasic signal conditions. Although no waveforms of basilar membrane displacements in basal regions are available in the present study it would seem reasonable to extrapolate from the waveforms of Fig. 7.10 to those of points further towards the base. On this basis the intervals between successive unipolar displacement peaks would become smaller as points further towards the base are considered. Without being very precise about the actual temporal features of important displacement waveforms and their points of origin, it is possible to suggest that in the antiphasic conditions, images based on a cross-comparison of corresponding basal points, would be predicted which would follow trajectories symmetrical about and close to ITD = 0. This condition has arisen for other images, as discussed above (Section 7.4.1), but it is proposed that here the image trajectories are even more closely positioned than in those cases - so close, in fact, that the result would be a split-image situation even more difficult to resolve and which would encourage the reporting of compromise judgements. In short, it is suggested that the high-pitched impulsive image arises from the cross-comparison of neural activity arising in unspecified portions of the basal region, and that the judged image centerings near ITD = 0 represent in one case (cophasic) the centering of the anticipated image, and in the other case (antiphasic) the centering of a compromise image, representing a balance of attention between the closely positioned symmetrically disposed elements of a split-image.

In conclusion, the three coexisting impulsive images which may be identified in experiments using unfiltered acoustic transients appear to be well described in terms of substantially independent cross-comparisons of neural impulses arising from three distinct and closely corresponding regions of the cochlea; the image of lowest pitch from the extreme apical regions (those maximally responsive to frequencies near 400 cps), the commonly perceived dominant image from regions maximally responsive to frequencies near 1500 cps, and the high-pitched impulsive image from the basal regions. The proposed working picture of the cross-comparison mechanism has proved to be the only one of the four suggested modes of cross-comparison able to readily accommodate the perception of multiple-images, and capable of consistent and realistically precise predictions of experimental results.

The important question remains, however, as to the factors which determine the relative importances of the multiple cross-comparisons such that some apparently contribute to the perception of images, while others do not. In the experiments just discussed (Table 7.5) it was clear that the dominant impulsive image was associated with the region of the basilar membrane which was presumably giving rise to the most significant neural information. In this case it was apparently a combination of the dynamic properties of the basilar membrane and the noise masking conditions which established the points of cross-comparison. However, there would seem to be no evidence, based on measurements of basilar membrane displacement amplitudes, that any other regions of the cochlea should be especially prominent in the cross-comparisons.

How, then, do the other two impulsive images arise?

At the present stage there would seem to be no outstandingly clear answer. It is possible that factors other than that of displacement amplitude are operative. On the other hand, it may be that the mechanism is sensitive to more subtle features of basilar membrane displacements than are revealed by the broad outlines of curves measured using a model.

In respect of the impulsive image of lowest pitch which was identified in the preceding experiments, there is indirect evidence which at least supports the proposition that it is associated with neural activity in the 400 cps region of the cochlea. Guttman and Flanagan (1964) who were involved in a study of pitch perception using repetitive acoustic transients, concluded from their investigations that, "a major source of the pitch cue of low-frequency unfiltered stimuli is roughly the region of the basilar membrane sensitive to 400 cps." It was argued that neural activity in the region of the 400 cps point may be given special weight because the amplitude of basilar membrane displacement falls relatively fast apically, thus causing an increase in the importance of neural information arising just short of the most apically responding point. It would seem possible that this apparent enhancement of neural activity in the 400 cps region may be associated with the lowest-pitched impulsive image identified here.

It is possible that the remoteness of the extreme apical region from the predominating medial (i.e., 1500 cps point) region is also a factor promoting the independent identification of low-pitched images associated with neural activity arising from this region. The high-pitched impulsive image may be attributable to similar factors which relate to the relevant (basal) regions of the cochlea.

In any event, it has been clear that these latter two images are of secondary importance; the majority of listeners did not recognize their presence. In fact, only highly experienced listeners have reported hearing any impulsive images other than the dominant one. It is therefore also possible that the experience of sustained critical listening under specialized experimental conditions has enabled these listeners to divert their attention from the dominant image to the more subtle images.

In the case of experiments which used filtered transients in one channel and unfiltered transients in the other (Figs. 7.4-7.7, Tables 7.3, 7.4 - Experiment B) it was found that the results were also best described in terms of a cross-comparison of neural information arising from corresponding regions of the cochleae. The point of cross-comparison was apparently associated with frequencies estimated to be near the cutoff frequency of the low-pass filter. Again, there is the question of why this region is of particular importance.

As stated earlier in the relevant discussion, it was clear that the low-pass filtered transient signal would produce a pronounced peak of basilar membrane displacement in the regions associated with the filter cutoff frequency (Fig. 7.14). However, there is a pronounced peak of displacement near the 1500 cps point of the contralateral basilar membrane (Fig. 7.11), which might also be demanding of attention. The difference in these two situations might be resolved in terms of the signal bandwidths; the wideband (unfiltered) signal could be associated with a generally broad peak of cochlear response, centered on the 1500 cps point, whereas the narrow-band (filtered) signal may be associated with a narrower peak of cochlear response, confined to the extreme apical regions. The last remark has been purposely stated in terms of 'cochlear response' because it is clear that not only factors of basilar membrane displacement are involved (although these may often be the dominant factors); the linearity characteristics of the neural transductions must also be taken into account. The discussion of Galambos and Davis (1943), by Allanson and Whitfield (1956) is relevant here. In Fig. 2 of Allanson and Whitfield are shown stylized curves of mean pulse rates in the cochlear nerve fibre array (arranged according to distance along the cochlea) which illustrate well the non-linear relation between mean pulse rate and stimulus intensity. From these curves it is apparent that for single frequency signals (and presumably narrow-band signals) the distribution (along the cochlea) of neural firings is sharply peaked and confined to a narrow region at low signal levels (i.e., small basilar membrane displacements), and becomes broader and flat-topped for higher signal levels.

Translating these results into terms of the present experiments it could be concluded that the smaller and narrower peak of basilar membrane displacement associated with the narrow band signal (e.g., Fig. 7.14) gives rise to a more sharply defined neural response along the cochlea, than does the larger proportioned peak of displacement associated with the wideband signal (Fig. 7.11). The consequence of the difference may be a favouring of the cross-comparison associated with the sharper peak of neural response. Naturally such an explanation is provisional because direct quantitative relations have not yet been established between rates of neural firing and amplitude of basilar membrane displacement. The final solution of this question must await further neurophysiological data.

In summary, it has been deduced that, of the four modes of cross-comparison suggested earlier, the experimental results of the present chapter are most satisfactorily described in terms of that which formed the basis of the proposed working picture of the mechanisms of cross-comparison. The multiple impulsive images reported in experiments using binaural unfiltered transients have been accommodated only by that hypothesis according to which multiple cross-comparisons are effected between neural impulses arising from corresponding narrow regions of both cochleae. The processes by which the multiple images are established, seem to be closely dependent on the peaks of basilar membrane displacement and the non-linear properties of mechanical-to-neural transductions.

It is important next to briefly consider a few experiments of earlier chapters with a view to appraising in a more general sense, the utility and applicability of the proposed working picture.

7.5.4 Discussion of Some Earlier Experimental Results in the Light of Present Views on the Mechanism of Cross-Comparison

Perhaps the clearest support to the proposed framework of cross-comparison developed out of work described in previous chapters has been that provided by experiments using selective masking procedures (Section 6.4.5). In those experiments it was seen that the application of filtered masking noise,to either or both channels, in no important way altered the lateralization characteristics of the remaining roving image; for example, see results illustrated in Figs. 6.27 through 6.31. Noise applied to one channel apparently prevented any important fusions of neural information arising from the masked regions of the cochlea. The result in such cases was an image having a pitch-like character appropriate to the band of masked frequencies, which was fixed at the extreme position on the unmasked side (the 'residue' image in earlier discussions). Similar noise applied to both channels apparently eliminated the image. In terms of the descriptive framework of Fig. 7.21(b), neural impulses from those regions of the cochlea which were activated by the random masking signal were rendered largely asynchronous, and hence in the relevant cross-comparisons did not produce a fused image of significant magnitude. The static images which were identified in the unilaterally masked conditions could, on this basis, be described as monaural images. On this view, lateralization properties of those images arising from cross-comparisons of neural activities from unmasked regions would be largely unaffected.

However, factors other than those being currently discussed may be important in the lateralization process; for example, an apparent exception to the latter statement was the experimental result depicted in Fig. 6.28 (described in Sections 6.4.5 and 6.5.0). There, the trajectory of the low-pitched S_2 image was clearly revealed in the judgement curves when the high-pitched impulsive image was rendered immovable by unilaterally applied high-pass filtered masking noise. However, it is believed that in this instance the masking operation merely eliminated a distracting influence (i.e., a clearly fused, high-pitched impulsive image), because on an earlier occasion (Fig. 6.7) the same behaviour was observed in the presence of the high-pitched image and without noise masking.

It will be noted in passing that the image-lateralization results of Fig. 6.31 are of the same form as the masking results reported in Table 7.3, which were obtained using the centering method.

A further matter of some importance is that relating to the monaural temporal masking mechanisms as discussed in Sections 6.2.1 to 6.2.4. It was there found (using repetitive pulse versus pulse pairs) that the low-pitched impulsive images apparently were not seriously influenced by monaural temporal masking effects, whereas the higher-pitched impulsive images were clearly and similarly affected. The differences were only clearly apparent at small δt values where the trajectory of the high-pitched S_2 image was usually not observed at all (e.g., Figs. 6.13a, 6.21, 6.24b), while that of the low-pitched impulsive image was readily apparent (e.g., Figs. 6.19a, 6.19b and 6.10).

In an attempt to cast some light on this problem the measurements of Figs. 7.15-7.18 were made using the basilar membrane model (see also Section 7.4.2). It was evident that for wideband transients at the smaller (less than 5 msec, say) pulse-pair spacings, the δt temporal information would be clearly maintained in basilar membrane responses down to some apical point which is maximally responsive to frequencies near $1/\delta t$ cps. At points more apicalward the important waveform features indicating δt appear to be rapidly diminished. At larger δt values the temporal information relating to the members of the pulse pair would be expected to appear in mechanical responses throughout the length of the cochlea. The important feature of these measurements is, however, that there is an enhancement of displacement amplitude in the region of the basilar membrane maximally responsive to frequencies near 1/δt cps; see Figs. 7.17 and 7.18 for comparison with Fig. 7.14. In short, when signals have spectral energies such that substantial basilar membrane displacements occur in the extreme apical regions, it is proposed that at the smaller δt values, an increased weight may be given to neural responses from the region of the cochlea selective to frequencies near 1/δt cps. For example, in the case of single- versus double-repetitive transients, where δt = 1.4 msec, the enhanced displacements would be expected at the 500 cps point (Fig. 7.16). On the working basis of cross-comparison, the image resulting from cross-comparison of neural information from these regions of both cochleae may be given increased importance; in particular, the neural concomitants of displacement waveforms shown in Fig. 7.10 (500 cps point) and Fig. 7.15 (500 cps point, δt = 1.4 msec) would be the important factors. It will be seen that on the basis of these waveforms the important image centerings would be anticipated at ITD's of about zero and 1.4 msec. Inspection of previous experimental results reveals that judged centerings for the low-pitched impulsive image were indeed spaced by closely δt in such cases (e.g., Figs. 6.10, 6.19).

If the important temporal information pertaining to the low-pitched impulsive image is provided by, or enhanced by, neural signals from the regions selective to 1/δt cps, the common interpretation of monaural masking (which considers only the acoustic signal - Guttman, et al., 1960) is not applicable. In this region (say, 500 cps) the auditory nerve fibers would be stimulated only once for each of the pulse pair members (because of the convenient spacing of successive displacement peaks). However, in more basal regions the membrane displacement characteristics are such that the associated nerve fibers would be repeatedly stimulated by displacements due to the first of the pair of transients (e.g., see Fig. 7.10), and presumably more of these fibers would be refractory on the occasion of the second transient of the pair; therefore, the neural response to this transient would be diminished. In this respect some results of Kiang, et al. (reported in Weiss, 1964) are of interest. PST (post stimulus time) histograms of single fiber responses to brief acoustic transients are shown which would seem to illustrate the remarks just stated.

The important conclusion is, therefore, that in considerations of monaural temporal masking, one must consider not only the temporal spacing of the successive acoustic stimulations, but also the dynamic properties of the basilar membrane and the images which may be associated with different regions of the cochlea. Signals having spectral energies extending to the very low frequencies (near $f=1/\delta t$) may apparently be expected to enable both S_1 and S_2 image lateralizations at smaller δt magnitudes than signals which have been filtered so as to exclude these frequencies; see results illustrated in Figs. 6.10 and 6.19 compared with those of Figs. 6.20 to 6.25. Equally important, where multiple impulsive images may be identified the degree to which monaural masking effects are reflected in the results is a function of the image which is being attended to, or which dominates the listener's impression (see Figs. 6.5, 6.6, 6.7).

Related to enhancement of basilar membrane displacements discussed above, are the curves of Fig. 7.19 which illustrate yet another form of modification of membrane responses. For a given point on the basilar membrane the successive responses to periodic wideband transients are seen to be largely independent for repetition frequencies much lower than that at which the membrane point is maximally responsive. As the pulse repetition rate was increased, the membrane displacements were seen to become almost smoothly oscillatory and then diminish as the pulse frequency was increased above the characteristic frequency of the membrane point.

Another important finding of earlier work related to the appearance of subsidiary image trajectories. In the results illustrated in Figs. 6.20 to 6.26, subsidiary image trajectories were observed at spacings from the main-image trajectories which were related to the higher cutoff frequency of the similar bandpass filters in each channel. It is believed that, in each case, these results reflect the temporal pattern of basilar membrane displacements in the maximally responding regions (which are associated with the cutoff frequency, Fig. 7.14). This would be in accord with the proposed working picture of cross-comparison, and is the view tentatively proposed at the end of Chapter 6.

In summary, it has been shown that the envelope of basilar membrane displacements is a function of a number of signal parameters, as follows:

(1) signal bandwidth - wideband transient signals generally give rise to a dominant peak of displacements in the region maximally responsive to frequencies near 1500 cps (Fig. 7.11). Low-pass filtered transient signals produce a form of response modified to show a displacement peak at a point associated with the low-pass cutoff frequency (Fig. 7.14).

(2) pulse-pair spacing - double transient signals give rise to an enhancement of membrane displacements in regions maximally responsive to about $1/\delta t$ cps, (Figs. 7.15-7.18).

(3) repetition period - repetitive transient signals cause an enhancement of responses in regions maximally responsive to about 1/T cps (Fig. 7.19).

It is believed that these characteristics of cochlear dynamics may be operative in the mechanism of cross-comparison by altering the overall weighting of neural responses from the various regions of the cochlea, and hence the relative importances of the various images which are hypothesized to arise.

7.5.5 Discussion of Some Relevant Publications

Straightforward comparisons of the results reported above with those of other workers is difficult. There are a number of reasons for this, but generally they relate to differences in experimental techniques employed by the various researchers.

It has been a major finding throughout this study that the reliable and independent identification and lateralization of acoustic images are key factors in resolving the complex acoustic-image phenomena. It has been shown that, in experiments using binaural acoustic transients, multiple images are the rule rather than the exception. These images, when lateralized, have often been shown to have quite different lateralization characteristics. It has also been pointed out that, in experiments where listeners were required to lateralize only the dominant image impression, the results indicated that the various images may, at different times, contribute differently to the lateralization judgements.

The results of experiments using the image-lateralization technique have been generally reproducible and amenable to interpretation. Used in conjunction with multiple-image lateralizations it has been shown to be a powerful instrument enabling interpretations of a number of otherwise ambiguous situations.

In contrast, there are only a few published examples of multiple-image recognitions, and even fewer in which patterns of image position versus ITD have been obtained. Generally, the reported results were obtained using a form of image centering or positional matching. The latter method, wherein listeners match the position of a 'target' image with that of a 'pointer' image gives results which are particularly difficult to interpret because the method presumes some important knowledge of the lateralization characteristics of the pointer image (this is clearly of particular importance for a pointer image arising from binaural transients). The technique of centering has been used on a few occasions in this study, and can be of valuable assistance in certain instances; however, once more it appears to be of the utmost importance that the listeners should attend to specific images. In the vast majority of published experiments, listeners apparently attended to an undifferentiated image.

With that introduction it is interesting to examine some relevant findings of other workers. In a few instances there are points which it is not judicious to compare in detail because of the incompatible nature of the experimental techniques.

A number of writers have reported experiments in which selectively filtered or selectively masked acoustic transient signals were used. In most cases the studies were very restricted and employed one or another of the 'centering' or 'matching' techniques.

Opinions of the various authors have differed in a number of respects as to, inter alia, the source of neural information contributing to the perceived images, the number of images perceived in various circumstances and, indeed, whether or not images could be perceived in certain instances. It is therefore important to consider a few of these observations in the light of the results and hypotheses reported here.

Writers often reported hearing multiple-images in experiments which utilized different acoustic transients in each earphone. Harris (1960) has provided what is, perhaps, the least ambiguous evidence of this. His experiments used binaural repetitive transient signals in which the pulses were alternately high-pass and low-pass filtered. Under these conditions listeners reported hearing high-pitched and low-pitched images which could be independently moved about according only to the ITD between contralateral high-and low-pass filtered pulses. This finding is in agreement with the views presented herein, and has been experimentally verified in a similar experiment, as well as in the more important and also more difficult type of experiment which used wideband transients only, viz., the independent tracking of low-and high-pitched impulsive images.

Harris (1960) also put forward an argument favouring independent neural pathways which would carry timing information pertaining to the high-pass and low-pass click images. Other writers, such as David, et al. (1958, 1959) and Leakey, et al. (1958) had earlier postulated a form of functional dichotomy in the frequency domain. David, et al. (1959) proposed a descriptive model for the processing of complex stimuli wherein the frequency range was dichotomous. No proposals were presented as to the detailed physiological processes, but it was suggested that the mechanism which effects a comparison between responses from the two ears may operate on the combined neural responses.

Other authors have presented a slightly different argument. Presumably supported by substantial physiological evidence Deatherage, et al., (1959) and Deatherage and Hirsh (1959) argued that the neural information from the basal turn of the cochlea provided the important part of the cochlear whole-nerve action potential, the N_1 response. They postulated that the latency of N_1 was the physiological correlate of perceived time of a brief acoustic signal, and used the results of their physiological measurements in the design of several psychophysical experiments. Results of these experiments were interpreted as confirming these authors' views. However, later findings (Deatherage, 1961) have shown the initial assumption to be an oversimplification; the author then, with certain reservations, aligned his views with those of Harris (1960).

The later experiments (Deatherage, 1961) were interesting in that the writer reported multiple impulsive images in experiments where the filtered bandwidths of the signals did not overlap. Attempts to reproduce these results here were unsuccessful. The present writer and six other listeners of varying experience attempted to centre images using binaural filtered transients with no common bandwidth. The experiments were arranged precisely as those described by Deatherage, even to a comparison of measured acoustic waveforms. Listeners regularly reported two impulsive images of different character at contralateral extreme positions, but could not, even by manipulation of IAD as well as ITD, bring either image to the subjective central position. A few listeners reported a sensation of motion away from the lateral extreme position, but the impression was very unstable and the motion only slight. No explanation seems obvious for these contradictory findings, except, perhaps, the differences in listeners. Where the signal bandwidths overlapped, listeners in the present experiments could readily lateralize a single image which had an acoustic character clearly related to the frequencies within the common bandwidth. Franssen (1964) reported a similar view summarized as follows: "an auditory sensation of direction can be obtained only when the signals presented to the two ears lie in the same band of frequencies."

In brief, the present study has failed to confirm certain of the experimental observations reported by Deatherage (1961) and, more important, it has been found difficult to reconcile certain neurophysiological implications of that work with the conclusions to which the present author has been led. It would seem that observation of whole-nerve action poetentials of the type mentioned above, and those more recently reported by Teas, et al., (1962) may be misleading. In particular, it seems that the binaural mechanism is sensitive to more subtle and more numerous constituents of cochlear nerve activity than may be revealed by simple observations of combined whole-nerve action potentials.

In the reported experiments that have come to the present writer's attention, no mention has been made of multiple impulsive images which were perceived to coexist in conditions of normal, unmasked, wideband transient stimulation. To a certain extent this would be expected, because with the binaural signal configurations most commonly used the low - and high-pitched impulsive images were observed to move together under the influence of ITD. Occasionally, however, certain situations revealed the rather different lateralization properties of these images. One such situation was the binaural single- versus double-pulse configuration which has been the subject of considerable study here.

With the background of experience acquired in these experiments it is interesting to speculate about the different findings of Guttman, et al. (1960) and Harris, et al. (1963) who used signals similar to those employed here. Harris, et al. reported S_2 image centerings at δt values as small as 0.5 msec, and found the repetition period not to be an important factor. On the other hand, Guttman, et al. reported S_2 image centerings at values of δt only as low as 3 msec, and found repetition period to be an important factor. It is now possible to suggest that these two findings are equally understandable, and that the difference lies mainly in the images to which the listeners attended. It seems likely that the listeners in the experiments of Guttman, et al. attended to the dominant impulsive image (such as was identified in the experiments of Chapter 4), while those of Harris, et al., for some reason, chose to seek out in more isolated fashion the low-pitched image which may not have been always dominant (for example that of Fig. 6.7).

Closely related in one sense are those experiments in which multiple centerings of an impulsive image have been observed where the signals were binaural single transients presented in cophasic and antiphasic fashions (e.g., Fig. 7.2 and Table 7.1). The psychophysical findings of Flanagan (1962) have already been briefly discussed in connection with related experiments in this series. It was seen that Flanagan's image-centering results were closely reproduced in the results listed in Table 7.1, but it seems that on the basis of the present findings that the picture presented by this author was incomplete. In particular, Flanagan made no mention of multiple images.

Another attempt to reproduce the image-centering results of Flanagan was that of Nordmark (1963). This writer, who used filtered instead of masked acoustic transients, reported image centerings that were closely ITD = 0 for both cophasic and antiphasic transients. However, it was observed that in the antiphasic condition, as ITD was gradually increased from zero, the image moved away from the centre position and then turned back momentarily before continuing towards the lateral extreme receiving the leading signal. Secondly, it was found that both antiphasic conditions produced similar effects. Both of these findings are in accord with the results of the present study; however, it is believed that in this case, as in certain experiments reported in the present chapter, the ITD = 0 centerings reported in the antiphasic conditions represent compromise judgements (Section 7.4.1), and that the images tracked at ITD values slightly greater than zero are the symmetrically disposed images which would be anticipated. Again, failure to achieve centering would appear to be a poor criterion on which to judge the absence of an image. The equivalence of image lateralizations using both antiphasic polarity signals has been clearly shown in the present study (e.g., Figs. 7.4-7.7).

Teas (1962) conducted a series of experiments using acoustic transients of different waveform and reported that, on certain occasions, images exhibited multiple centerings. One technique used by Teas was closely related to the image-lateralization procedure employed here. In it he required the listeners to adjust the ITD until the image was coincident with centre, the eye (sic), the ear and positions intermediate to these. The results were, in some respects, similar to those reported in the present study, but before examining the results it would be well to correct a misleading label employed by that author. The two transient signals used in Teas' experiments were identified as "high-pitched" and"low-pitched." The former is really a misnomer because, in fact, the signal was wideband. The latter was more appropriately named because the signal was 600 cps low-pass filtered. With that in mind the result may be quite simply interpreted. It has been shown that a wideband transient apparently produces a peak of activity in the 1000-1500 cps region of the basilar membrane and that, on some occasions, this is believed to underlie the lateralization of a subsidiary image which follows a trajectory spaced about 1 msec from the main image trajectory (e.g., Fig. 4.29). Unless special procedures are invoked the appearance of this trajectory in the judgement curves is not consistent between listeners and between different experiments; however, where its presence has been inferred, the judgement curve often showed only a slight indentation after the main image had moved from centre to the extreme position. Such would appear to be the case in Fig. 2 of Teas (1962) for cophasic wideband transients.

When the bandwidth of the signal has been restricted by, for example, low-pass filtering the subsidiary image trajectories have been seen to be spaced at distances from the main image trajectory which are a function of the cutoff frequency of the filter (Figs. 6.18-6.25).

Teas also noted this phenomenon and curves of his Fig. 3 are here interpreted as reflecting the lateralizations of images arising from neural activities in the maximally responding (about 600 cps) apical regions of the cochleae.

Although absolute signal level was not a major parameter in the present study, casual experiments could not confirm the differences in slopes of the judged image trajectories reported by Teas (1962) as being a function of sensation level.

In the extensive works of von Békésy, it is not surprising that there are further findings of interest to the present study. Apart from the extremely valuable measurements of cochlear mechanics (which underlie the basilar membrane model) this worker performed some important pioneering experiments in the tracking of auditory images under masked conditions. From results of these experiments, which used single tones to mask neural activity in various regions, he concluded that "the direction of sound is not determined by the nerve activity of the entire length of the basilar membrane of both ears, but arises from the interaction of very narrow regions of the two membranes" (von Békésy, 1933 - reported in von Békésy, 1960, p. 318). This is clearly the basis of the working picture of cross-comparisons illustrated in Fig. 7.21(b).

On the subject of masking, there would seem to be fairly universal agreement that, if the masking signal is applied at sufficiently high levels, neural activity synchronized with the masked signal will be effectively eliminated. However, if the masking signal is at a level less than sufficient to dominate neural activity in the relevant cochlear region, there would seem to be some uncertainty as to its effect.

For instance, at the end of Section 6.4.5 an experiment was reported wherein unfiltered noise was unilaterally added to an unfiltered binaural transient signal at levels ranging over 20 dB, but in no case sufficient to completely mask the signal. The result was that neither the low-pitched impulsive image trajectories nor the second harmonic tonal image trajectories were appreciably altered; however, the high-pitched impulsive image was relegated to the unmasked extreme position at relatively low masking signal levels. No explanation was offered for this curious finding which suggests that so long as the signal is not entirely masked by the noise the roving image which is perceived is lateralized according to the signal IAD/ITD parameters and not the IAD measured as a relative interaural signal-to-noise ratio.

In a somewhat different context, David, et al. (1959) reported a similar observation. Their experiments, which were directed to a measure of the intensity-to-time conversion, indicated magnitudes of the trading relation which were, for the signals employed, a function of signal sensation level, and independent of binaural masking noise level. In a different experiment these same writers observed no change in the ITD required for centered images when the masking noise was applied at different levels in each channel.

Raab and Osman (1962) investigated this problem in some detail using a centering procedure wherein listeners adjusted the level of unmasked transients in one ear to offset the apparent IAD introduced by broadband masking of transients in the opposite ear. Findings suggested that the masking noise did cause a lateral displacement of the acoustic image away from the masked side.

Butler and Naunton (1962), in image localization experiments, observed that unilateral masking noise biased the positional judgements of the spatially-perceived image towards the side of the masking noise.

Clearly there is more work needed on this topic before an explanation is possible. It was because of the many inconclusive views on the use of masking noise at lower levels that the present study has used masking signals at levels sufficient to completely prevent detection of the acoustic transient signal.

It is relevant to the general arguments presented herein to consider a few results which used signals of non-transient nature. For example, some findings of Cherry and Sayers (1956) and Sayers and Cherry (1957) are of particular interest. These authors reported experiments in which intoned vowel sounds were presented binaurally with and without unilateral wide-band masking noise. Listeners in these experiments were required to report only dichotomous 'left' or 'right' judgements, indicating an important sound image to the left or right of subjective centre. With unmasked binaural signals the coherence curves reported by these authors indicated only one important centre crossing as ITD was altered through zero. When masking noise was applied at certain levels periodicities were observed in the patterns of sidedness judgements. These periodicities in ITD were found to agree closely with periods of the highest-energy harmonics as measured in a spectral analysis of the signal. It was therefore argued that the images being judged were the result of binaural fusions of the dominant harmonics of the binaural speech signal, the less important components having presumably been masked. Similar results were obtained in experiments using running speech

In this instance, periodicities of the coherence curves were observed to be different for male and female speech (longer period for male speech), presumably corresponding to the different dominant spectral components of each. Low-pass filtering of the unmasked speech signal was found to cause no alteration in the periodicity of the coherence curve unless the attenuation band of the filter encompassed a frequency corresponding to the periodicity in the coherence curve for that particular speech sample. When this was the case no periodicity at all could be seen (Sayers & Cherry, 1957).

Each of the above results are interpreted as supporting some general views of the present study. In particular, it seems that an enhancement of basilar membrane displacements by, perhaps, a prominent spectral component of a binaural signal may cause neural activity arising from the enhanced region to be of increased importance in the formation of images. There would be conditions where the independent identification of this image may be difficult or impossible because of distracting factors (i.e. perhaps other images), but there may also be circumstances, such as certain masked conditions, where this image is more clearly revealed. In these circumstances the listener's impressions may be dominated by this image, and the image lateralizations or sidedness judgements would reflect the periodicity of basilar membrane displacements in the newly-enhanced region and by inference, the approximate frequency of the significant spectral component.

7.5.6 Notes on a Recent Publication

Recent work by Flanagan, et al. (1964) is of direct relevance to the general conclusions of the present study. Much of this work was an extension of earlier image centering experiments using binaural unfiltered transients which have been masked with symmetrically filtered noise (Flanagan, 1962). As discussed above, results of the present study are in general agreement with these findings (Table 7.1), but it was noted that while Flanagan, et al. have reported multiple centerings of images, they have not reported multiple-image centerings; the distinction is an important one.

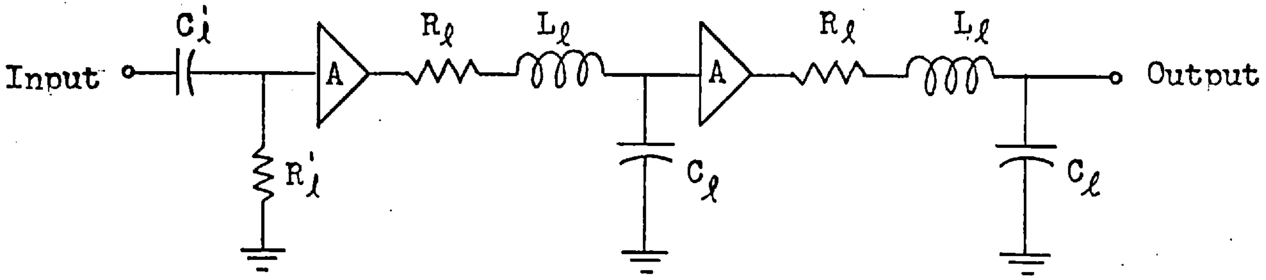
An important hypothesis put forward by these authors relates to the origins in the cochlea of the most significant neural activity. In particular, it was proposed that the important binaural images arise from the cross-comparison of neural activity from maximally responding unmasked regions of the cochleae. The authors used this hypothesis and the computed basilar membrane displacement waveforms (presumably in a manner similar to that employed in the present study) to predict image-centering values of ITD for binaural transient signals masked asymmetrically. For the single cophasic and single antiphasic conditions chosen by the authors, it was found that the predicted results were not always in agreement with the experimental results. It was proposed that the errors could be attributed to the erroneous assumption that maximum neural firings were associated with peaks of basilar membrane displacement.

The present study has confirmed the importance of images associated with neural activity apparently arising from maximally responding unmasked regions of the cochlea. However, the investigations of the present study into the manner of cross-comparisons have failed to show any consistent explanation for the experimental results other than that based on a cross-comparison of neural activity arising from corresponding regions of the cochleae. Many examples have been discussed (Figs. 7.3-7.7, Tables 7.2-7.5) which demonstrate for a variety of signals, the similarities of image lateralizations and image centerings in both cophasic and both antiphasic signal configurations. Predictive results based on cross-comparisons between maximally responding unmasked regions of the cochlea (as proposed by Flanagan, et al.) do not show such similarities, neither does such a view accommodate simultaneously perceived multiple-images.

The discussion by Flanagan, et al. (1964) of experiments using single- versus double-pulse trains is of further interest. The 'neural-gate' hypothesis proposed by Harris, et al. (1963) was again put forward as an explanation for an irregularity in experimental results. This topic was discussed in Chapter 4 above, where it was argued that there is no necessity for such a mechanism. There seems no reason to alter this statement.

FIGURE 7.1

Simplified Electrical Network Representation of the Model for Approximating Basilar Membrane Displacement



A = Amplifier Gain = 1

f_ℓ = Frequency at which the Membrane Point Responds Maximally

f_ℓ (cps)	C'_ℓ (μF)	R'_ℓ (K)	R_ℓ (K)	L_ℓ (H)	C_ℓ (μF)	Peak Amplitude Correction Factor (dB)
100	0.159	10	0.63	1.0	2.02	0
200	0.080	10	1.2	1.0	0.505	5.0
300	0.053	10	1.9	1.0	0.225	8.0
400	0.040	10	2.5	1.0	0.126	10.0
500	0.032	10	3.1	1.0	0.081	11.6
600	0.027	10	3.8	1.0	0.056	13.0
700	0.023	10	4.4	1.0	0.041	14.0
800	0.020	10	5.0	1.0	0.032	15.0
900	0.018	10	5.7	1.0	0.025	16.0
1000	0.016	10	6.3	1.0	0.020	16.7
1100	0.014	10	6.9	1.0	0.017	16.7
1200	0.013	10	7.6	1.0	0.014	16.7
1300	0.012	10	8.2	1.0	0.012	16.7
1400	0.011	10	8.8	1.0	0.010	16.7
1500	0.010	10	9.4	1.0	0.009	16.7

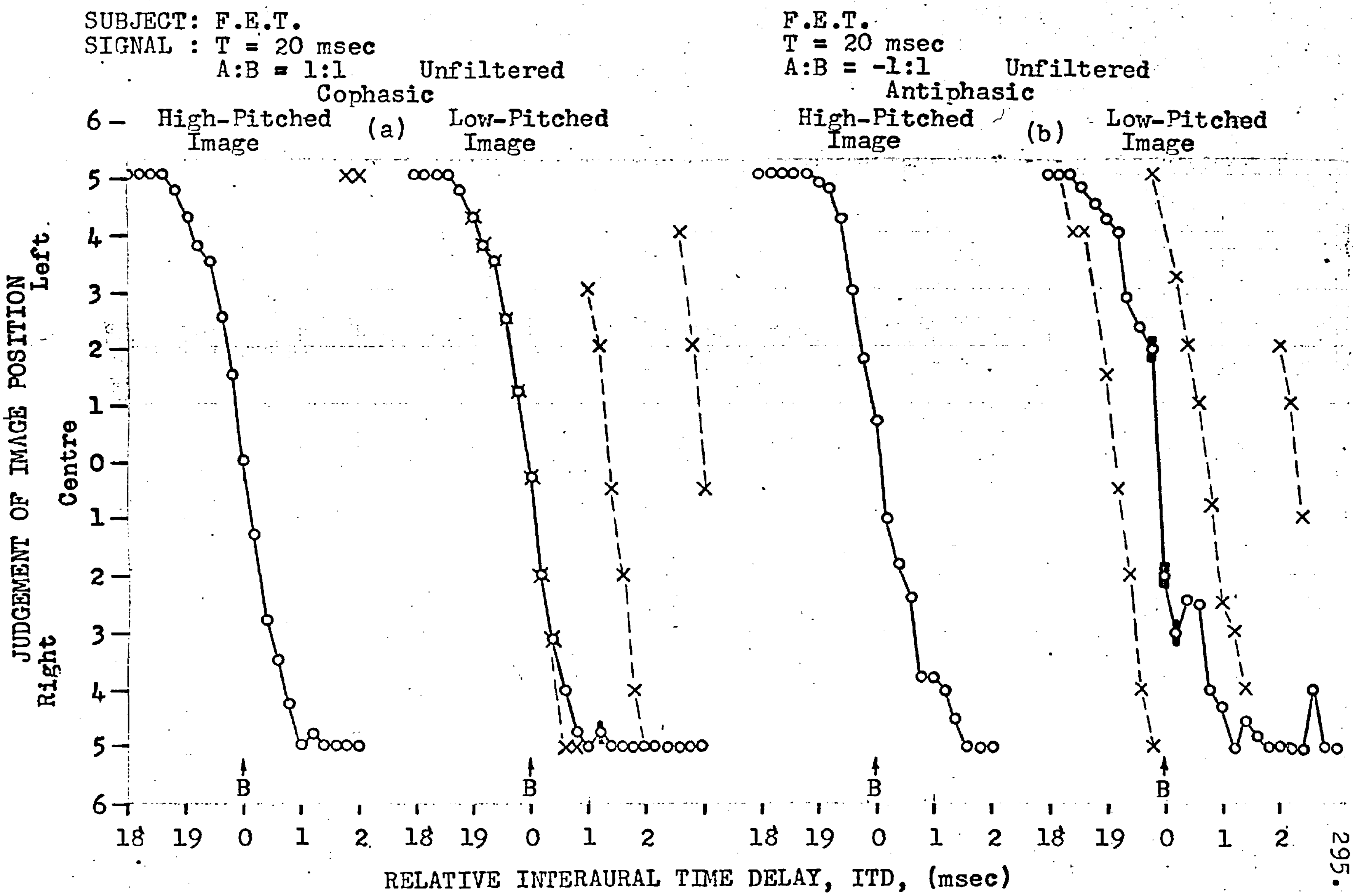
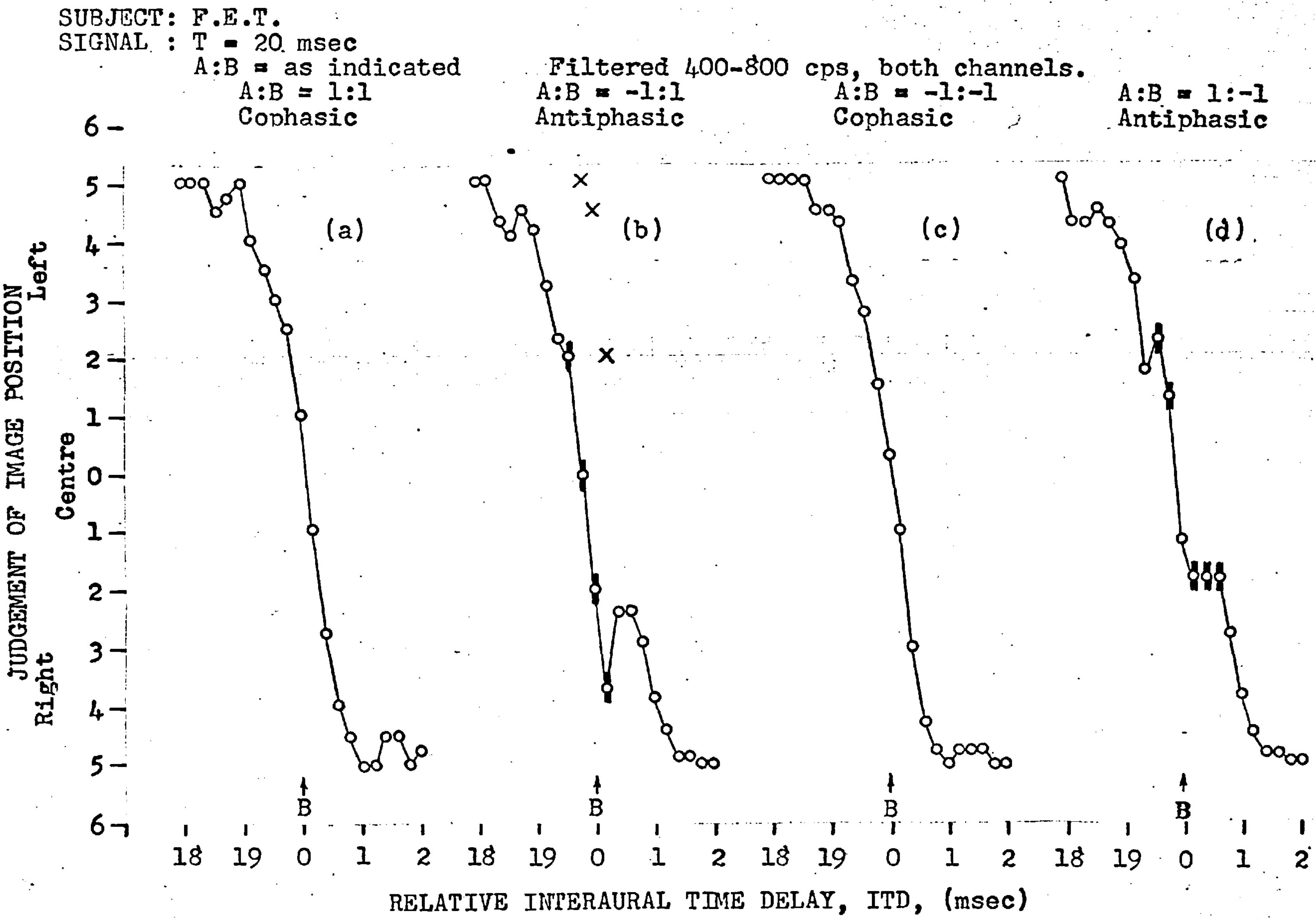

SUBJECT: F.E.T.
SIGNAL : T = 20 msec
A:B = 1:1 Unfiltered
Cophasic
High-Pitched Image
(a)
Low-Pitched Image
F.E.T.
T = 20 msec
A:B = -1:1 Unfiltered
Antiphasic
High-Pitched Image
(b)
Low-Pitched Image
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
18
19
0
1
2
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 7.2



SUBJECT: F.E.T.
SIGNAL : T = 20 msec
A:B = as indicated
Filtered 400-800 cps, both channels.

FIGURE 7.3

SUBJECT: F.E.T.
SIGNAL : T = 20 msec
A:B = as indicated

Bandpass Filtered 20-400 cps, Right Channel Only.

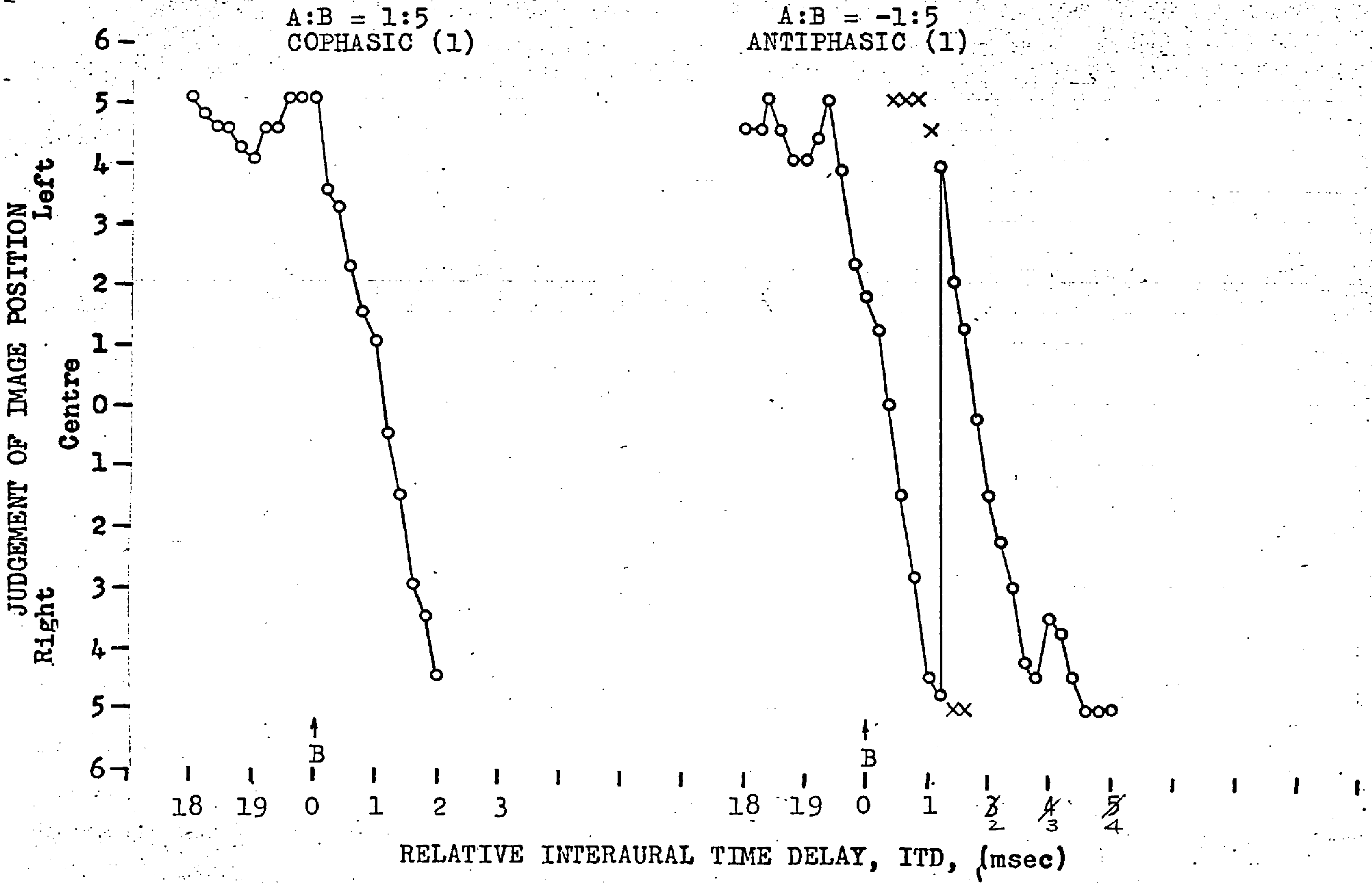


FIGURE 7.4

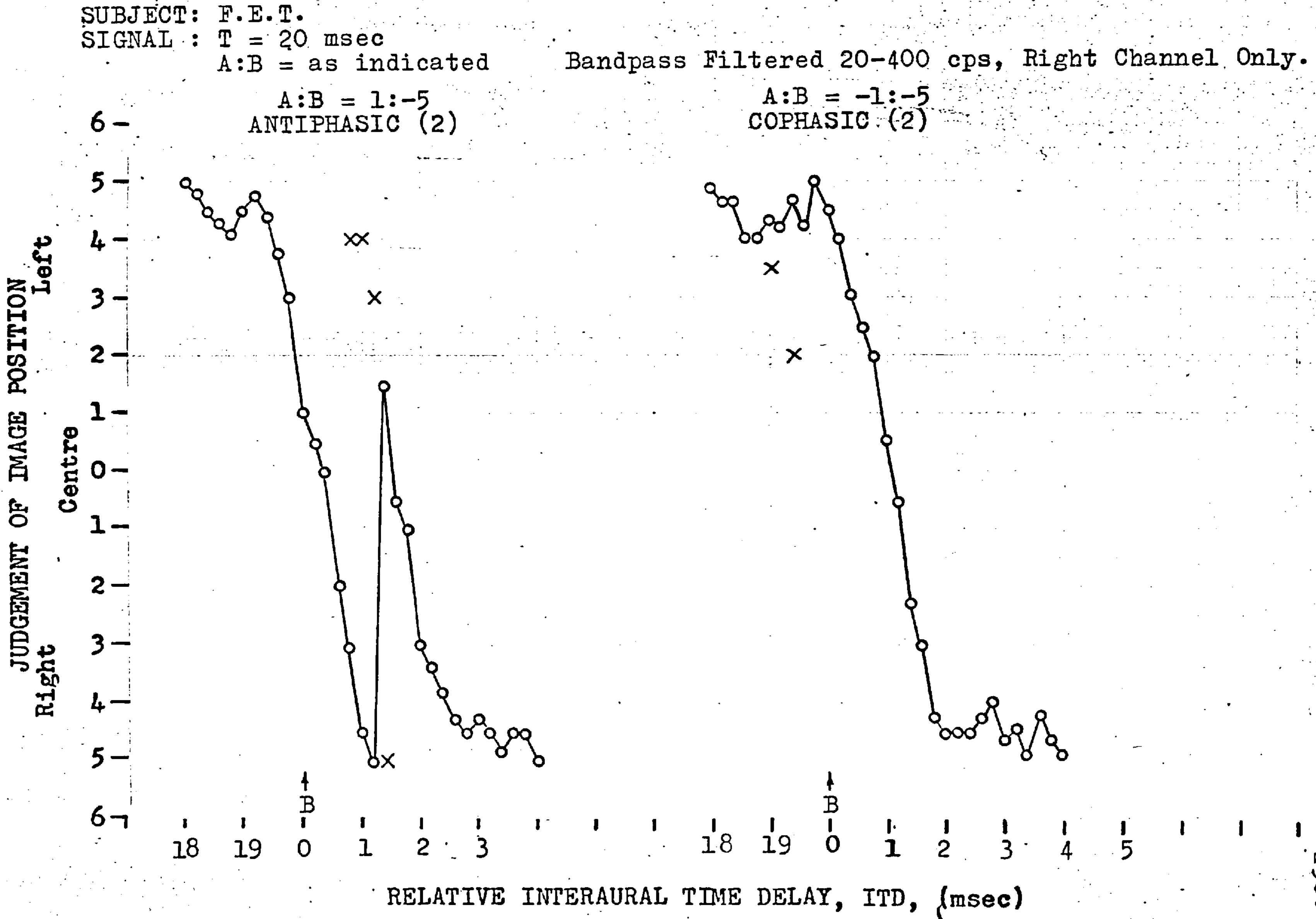
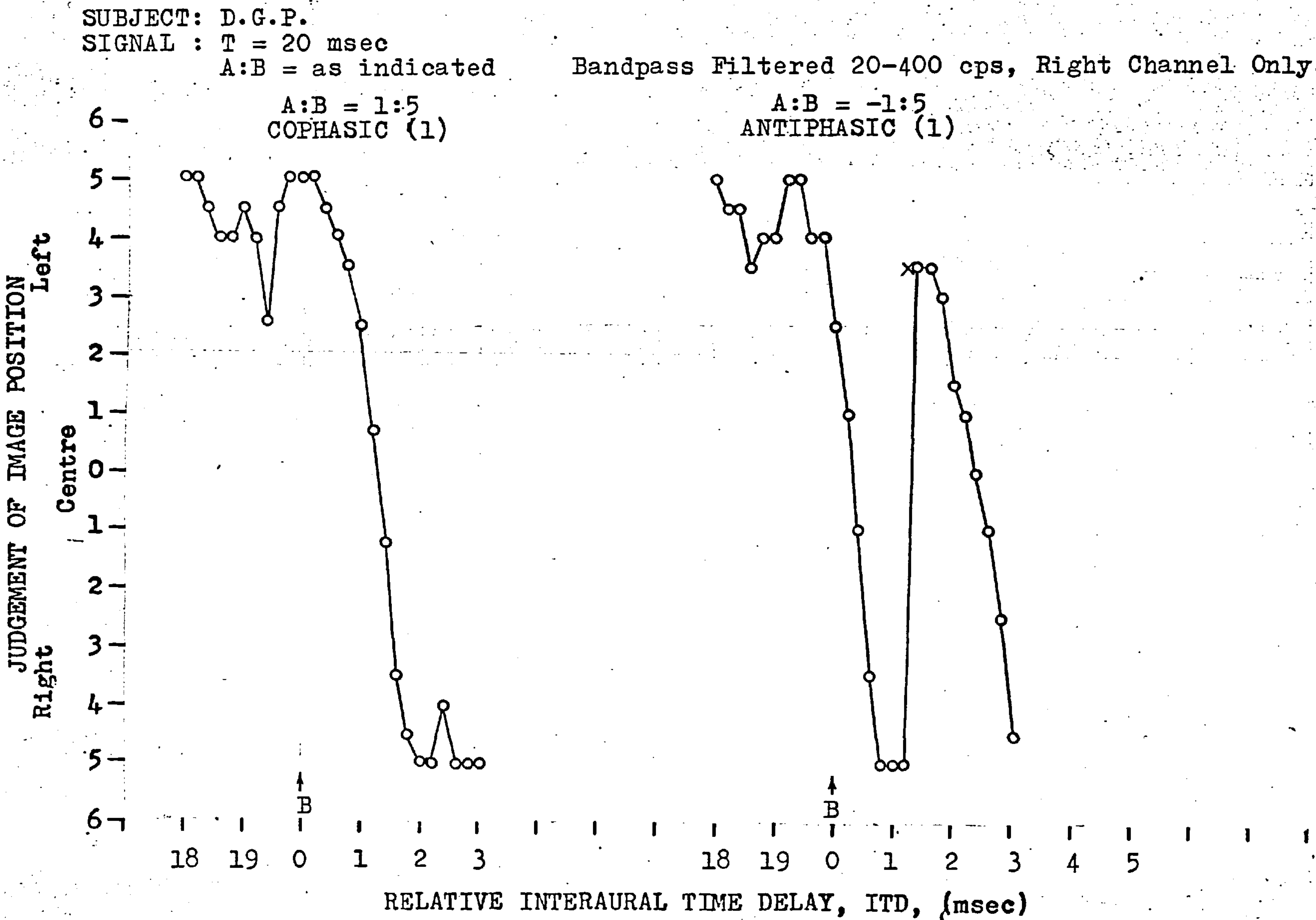

SUBJECT: F.E.T.
SIGNAL : T = 20 msec
A:B = as indicated
Bandpass Filtered 20-400 cps, Right Channel Only.
A:B = 1:-5
ANTIPHASIC (2)
A:B = -1:-5
COPHASIC (2)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
18
19
0
1
2
3
18
19
0
1
2
3
4
5
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 7.5



SUBJECT: D.G.P.
SIGNAL : T = 20 msec
A:B = as indicated

Bandpass Filtered 20-400 cps, Right Channel Only.

A:B = 1:5
COPHASIC (1)

A:B = -1:5
ANTIPHASIC (1)

FIGURE 7.6

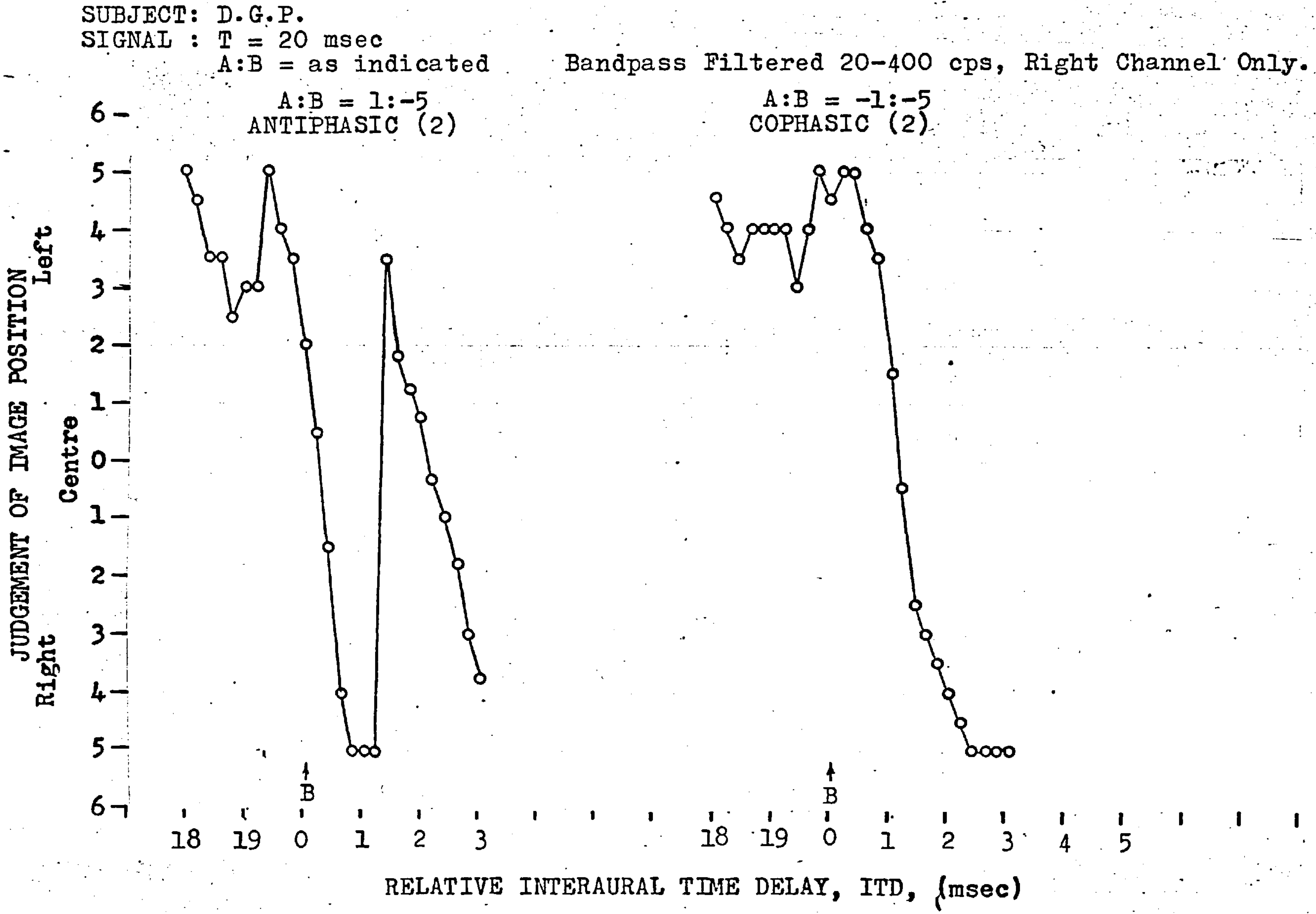

SUBJECT: D.G.P.
SIGNAL : T = 20 msec
A:B = as indicated
Bandpass Filtered 20-400 cps, Right Channel Only.
A:B = 1:-5
ANTIPHASIC (2)
A:B = -1:-5
COPHASIC (2)
JUDGEMENT OF IMAGE POSITION
Left
Centre
Right
6
5
4
3
2
1
0
1
2
3
4
5
6
B
B
18
19
0
1
2
3
18
19
0
1
2
3
4
5
RELATIVE INTERAURAL TIME DELAY, ITD, (msec)

FIGURE 7.7

SUBJECT: F.E.T.
SIGNAL : T = 20 msec
A:B = -1:1 Unfiltered
MASKING NOISE: As indicated, both channels.

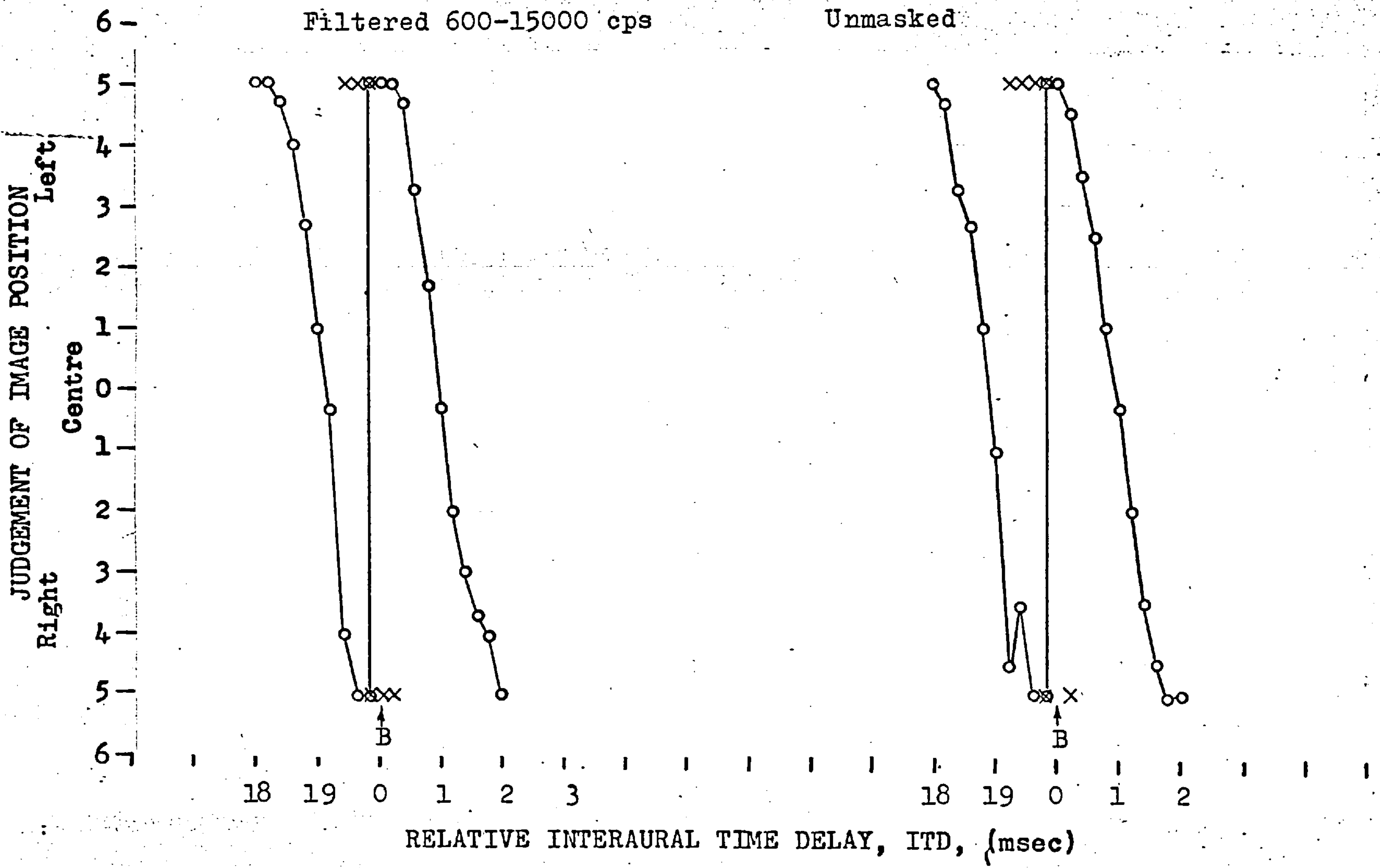


FIGURE 7.8

SUBJECT: F.E.T.
SIGNAL : T = 20 msec
A:B = -1:1 Unfiltered
MASKING NOISE: Both Channels, Bandpass Filtered 1000-15000 cps.

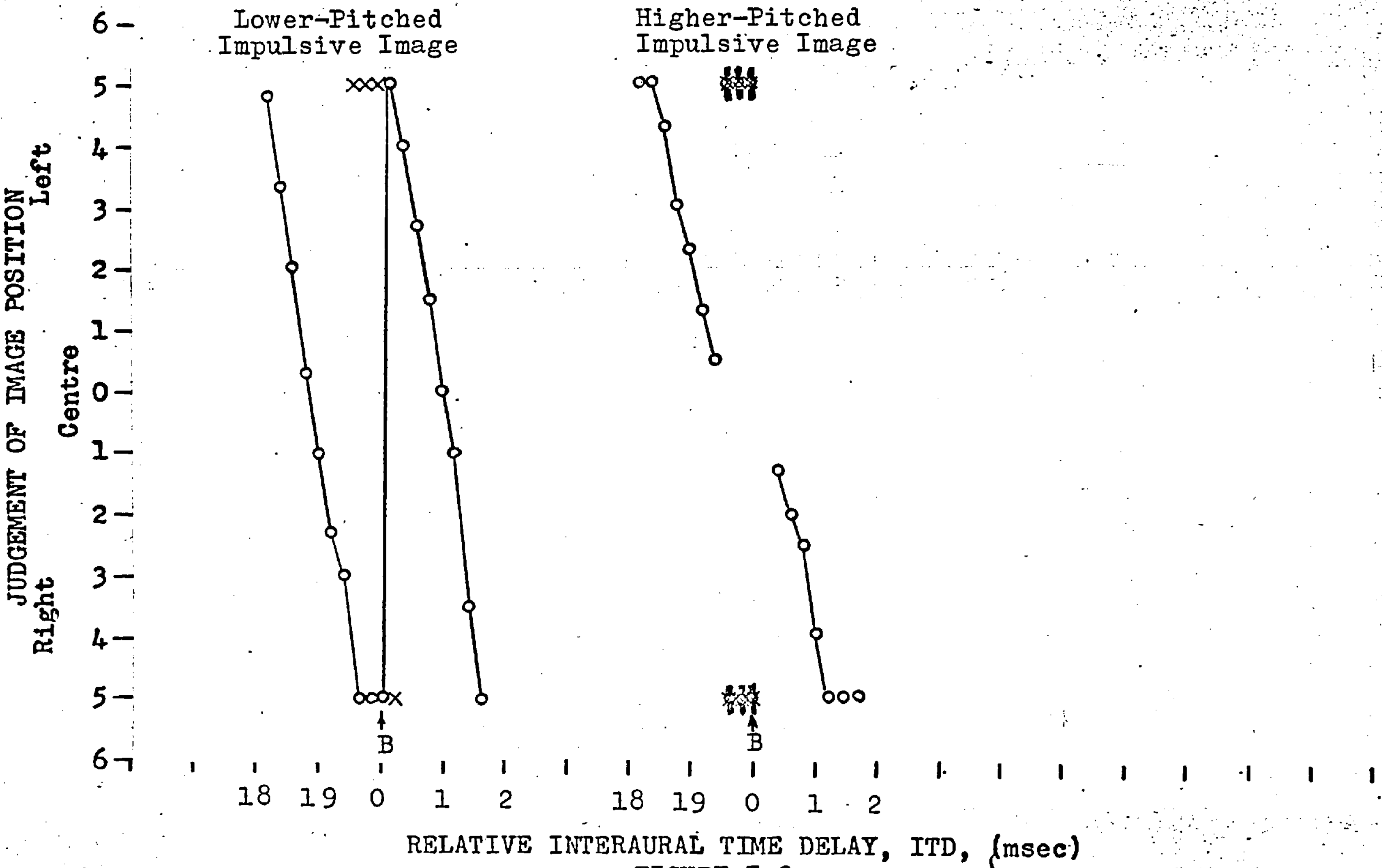
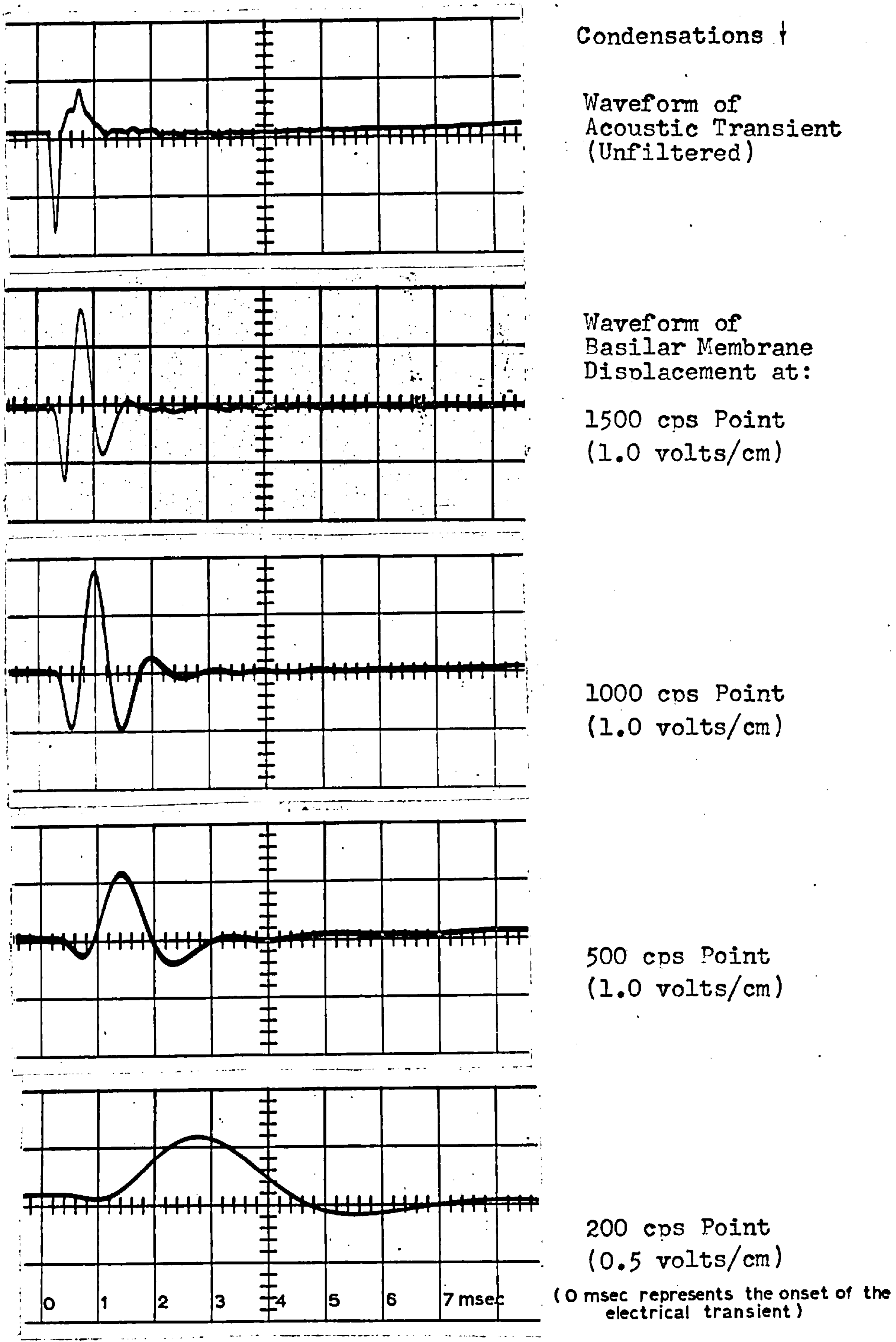


FIGURE 7.9

Waveforms of Basilar Membrane Displacements in Response to Unfiltered Acoustic Transients

FIG.7.10



(0 msec represents the onset of the electrical transient)

FIGURE 7.11

Corrected Basilar Membrane Displacement Amplitudes
Signal: Unfiltered Transients.

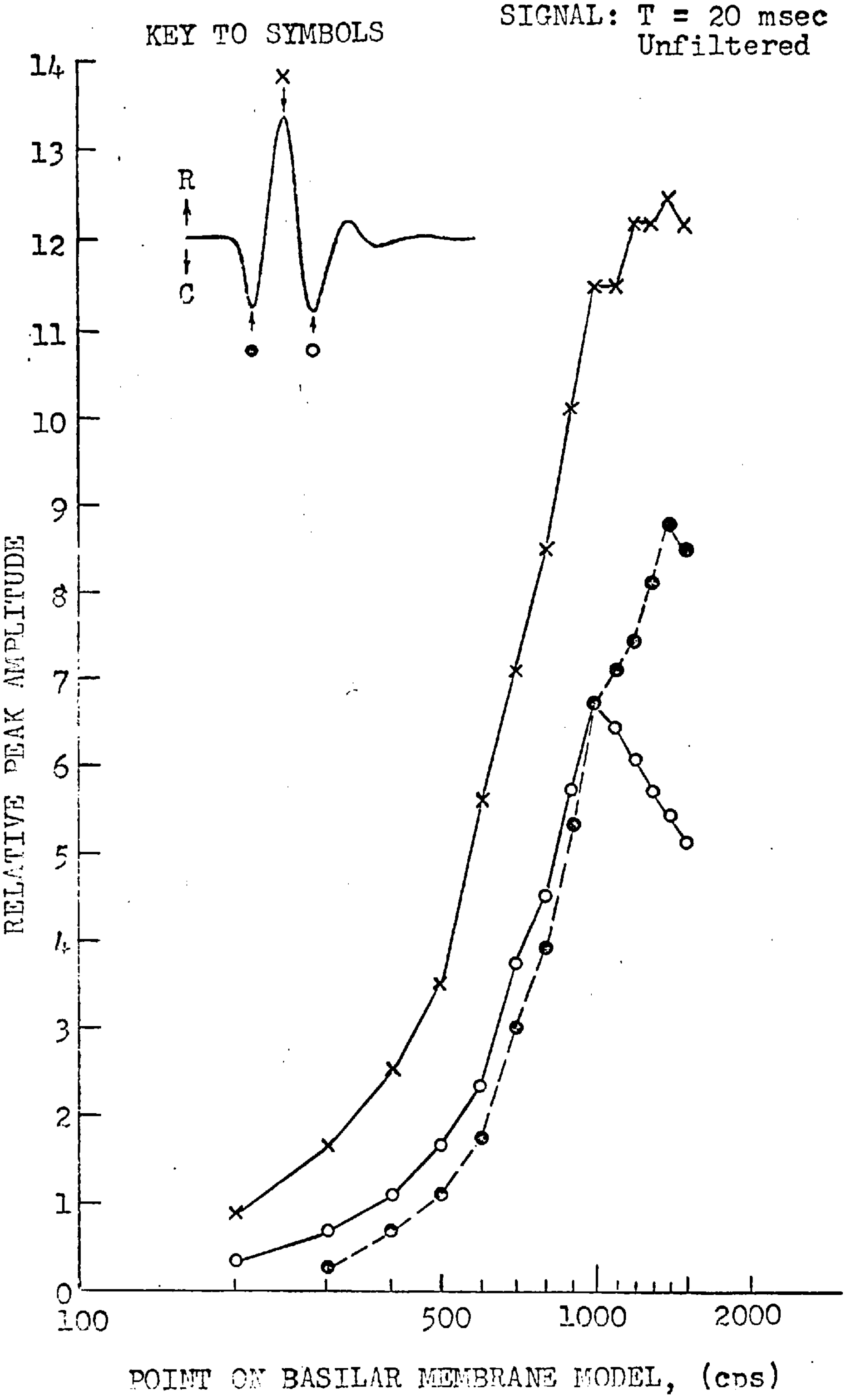


FIGURE 7.12

Waveforms of Basilar Membrane Displacements in Response to Filtered Acoustic Transients

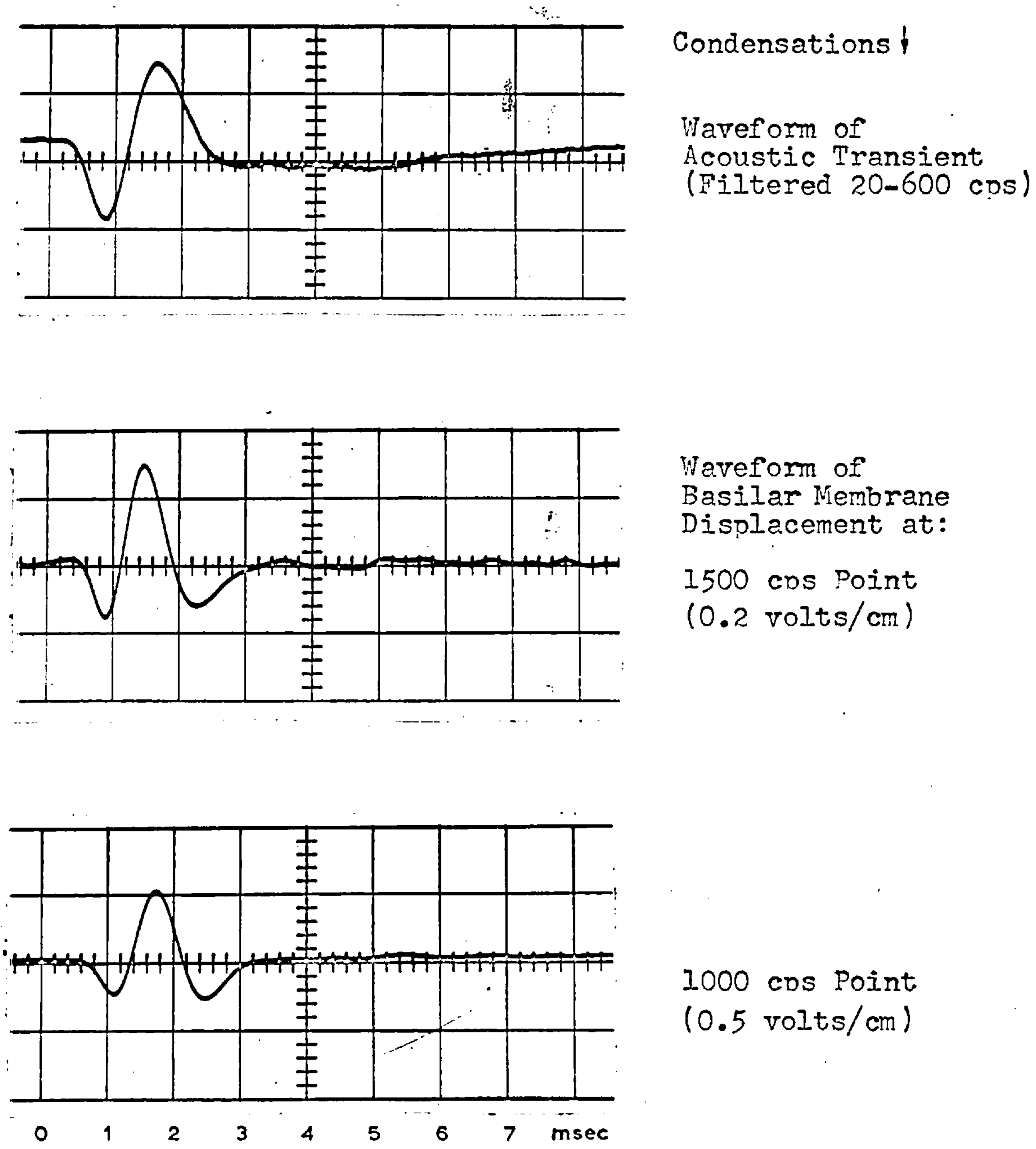


FIGURE 7.13

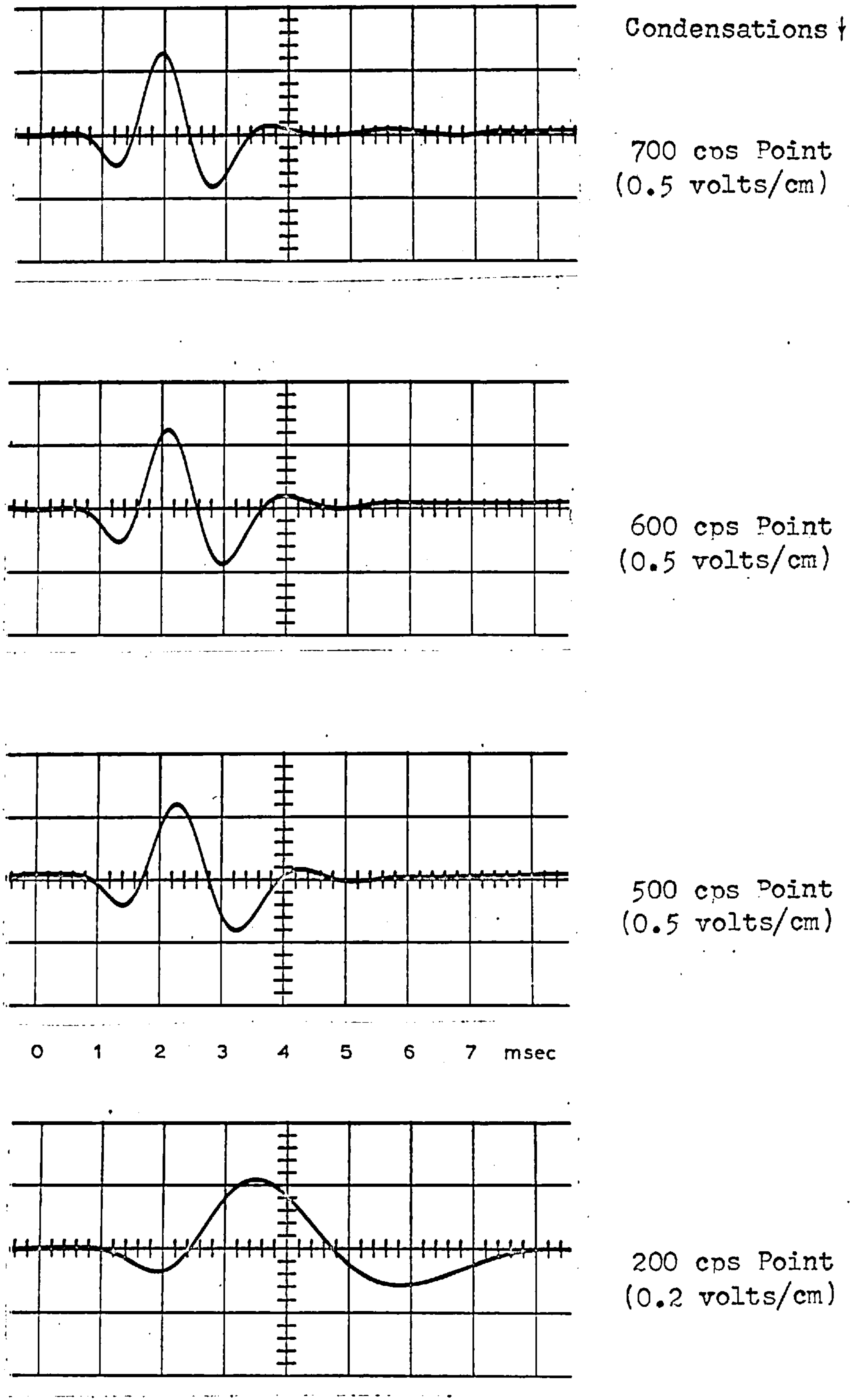
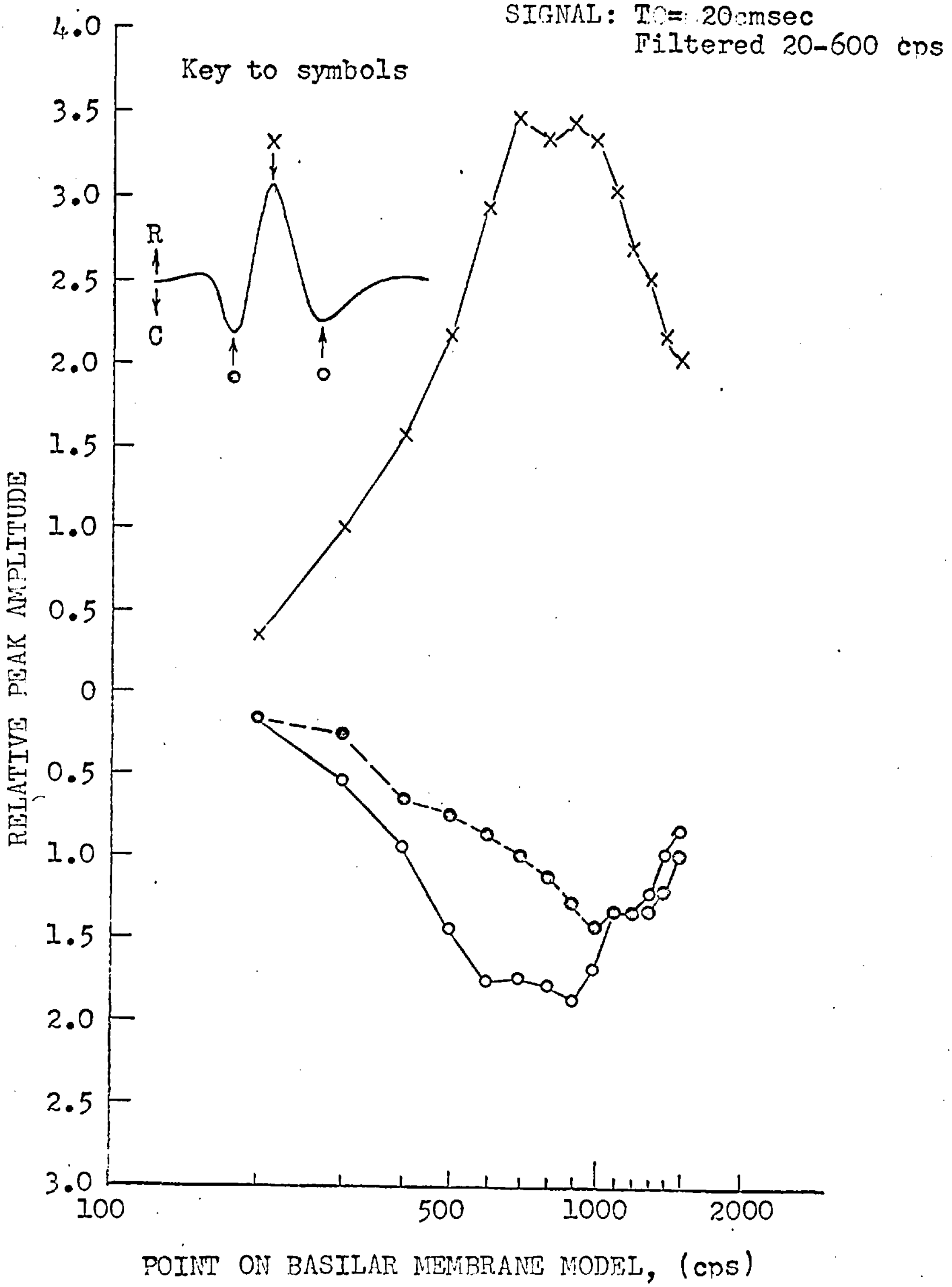

Condensations ↓
700 cps Point
(0.5 volts/cm)
600 cps Point
(0.5 volts/cm)
500 cps Point
(0.5 volts/cm)
0
1
2
3
4
5
6
7
msec
200 cps Point
(0.2 volts/cm)

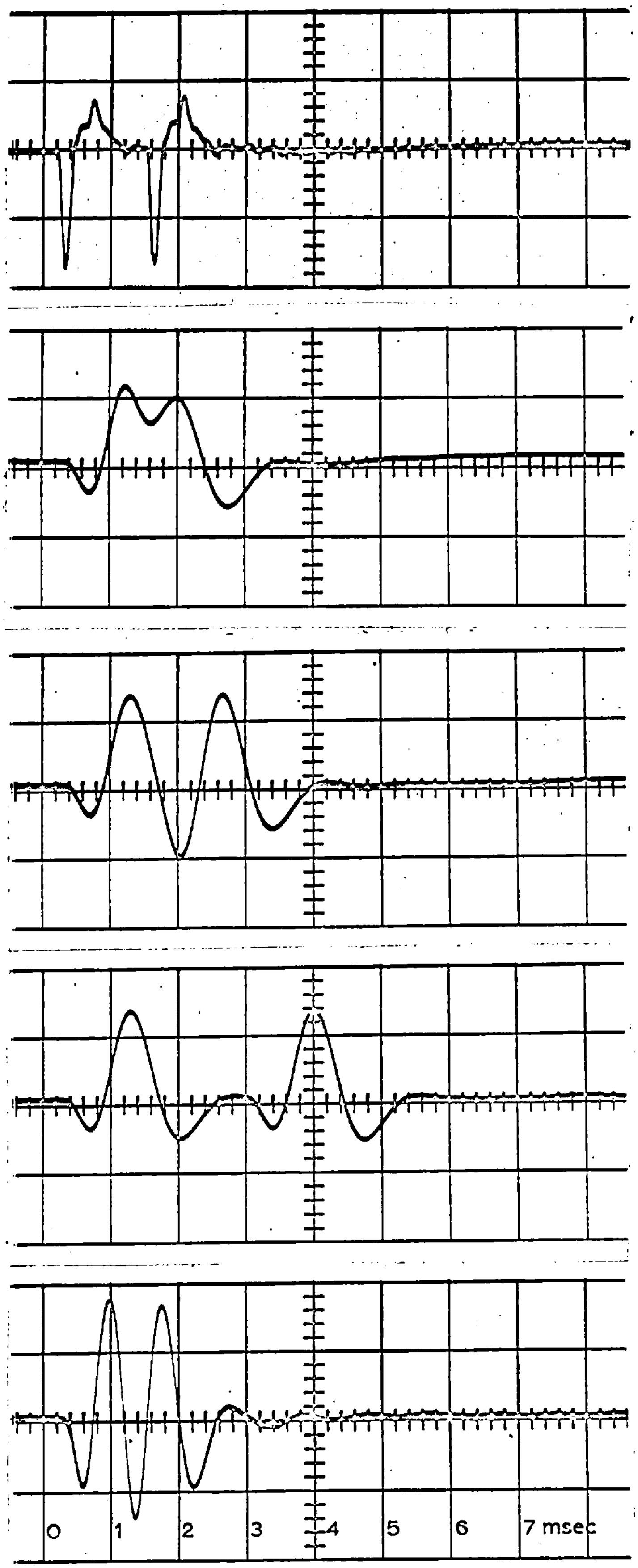
FIGURE 7.14

Corrected Basilar Membrane Displacement Amplitudes



The Effect of Pulse-Pair Spacing (δt) on Basilar Membrane Response

FIG. 7.15



Condensations ↑

Waveform of
Acoustic Transient
Pair (Unfiltered)
δt = 1.4 msec
(a)

Waveform of
Basilar Membrane
Displacement at:

500 cps Point
δt = 0.8 msec
(b)

500 cps Point
δt = 1.4 msec
(c)

500 cps Point
δt = 2.8 msec
(d)

1000 cps Point
δt = 0.8 msec
(e)

FIGURE 7.16

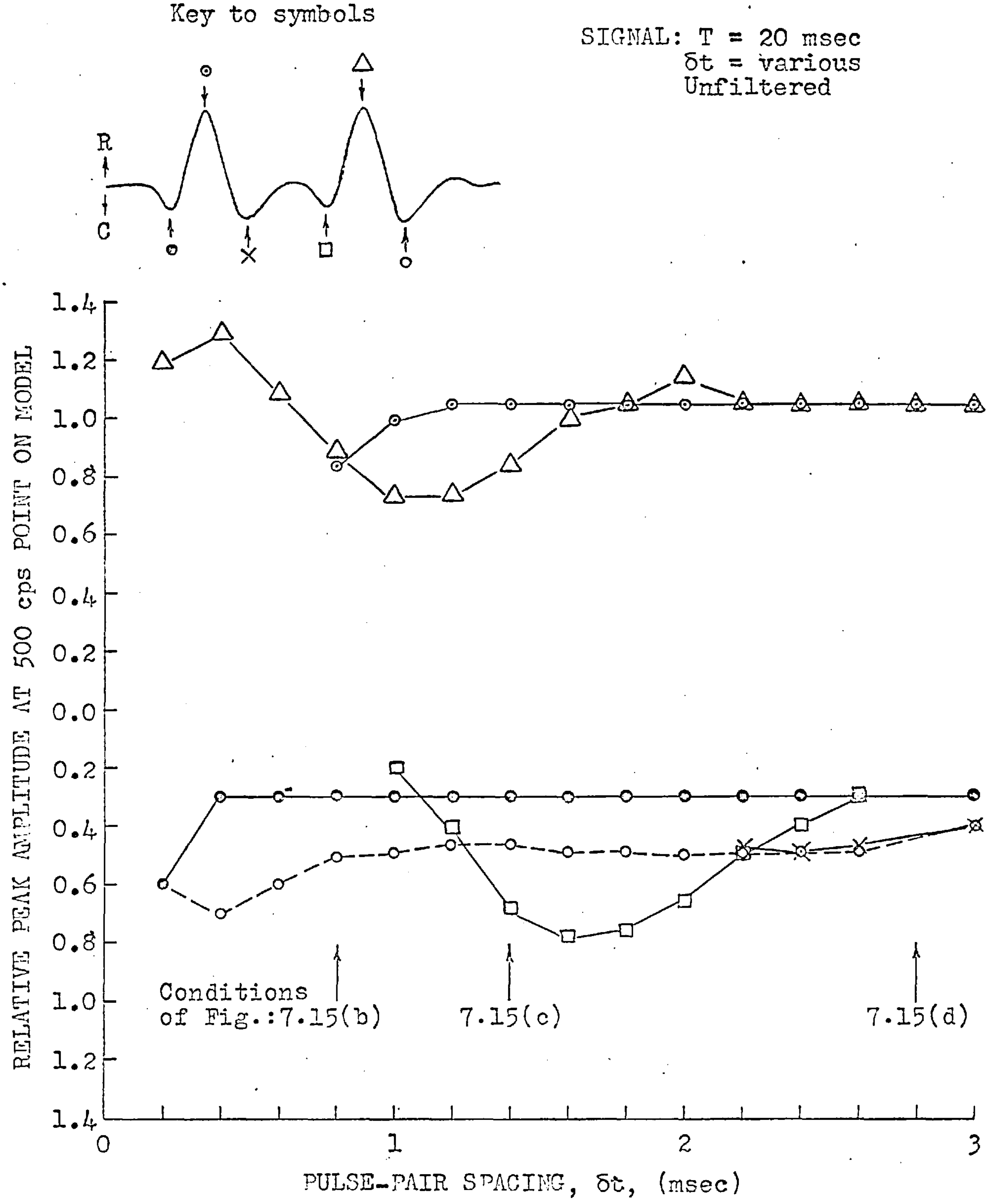

Key to symbols
R
C
SIGNAL: T = 20 msec
δt = various
Unfiltered
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
RELATIVE PEAK AMPLITUDE AT 500 cps POINT ON MODEL
Conditions
of Fig.: 7.15(b)
7.15(c)
7.15(d)
0
1
2
3
PULSE-PAIR SPACING, δt, (msec)

FIGURE 7.17

Corrected Basilar Membrane Displacement Amplitudes

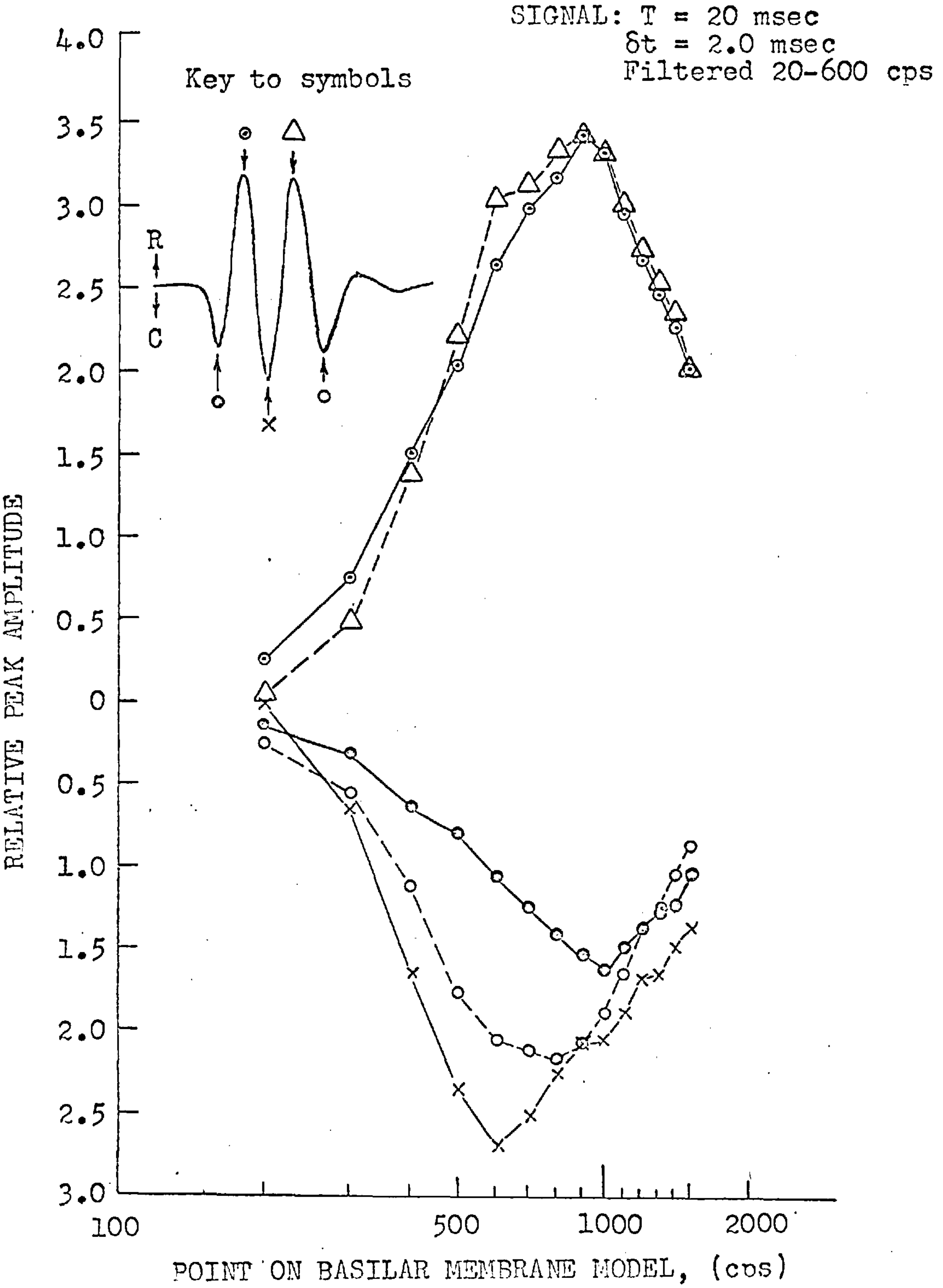


FIGURE 7.18

Corrected Basilar Membrane Displacement Amplitudes

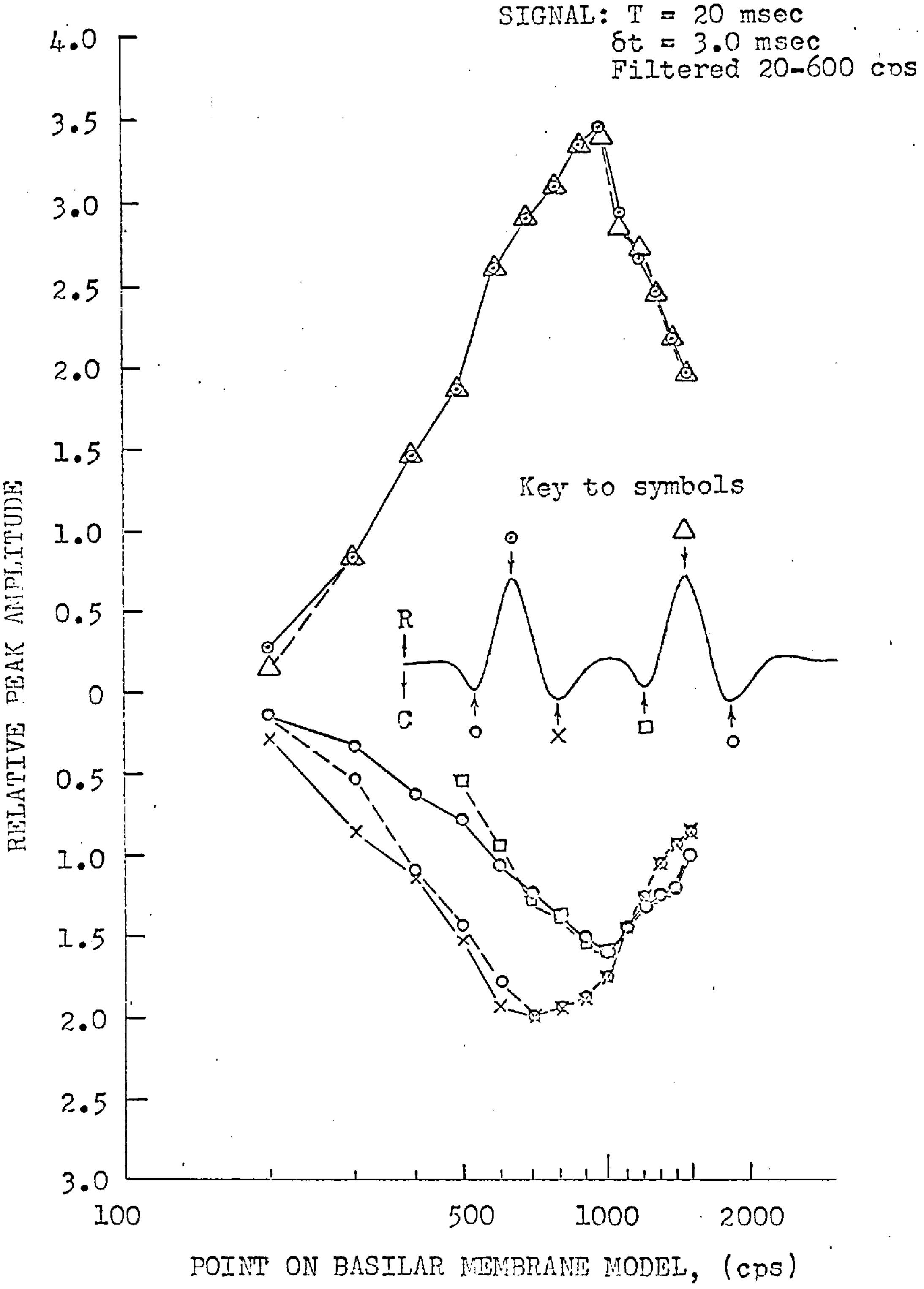
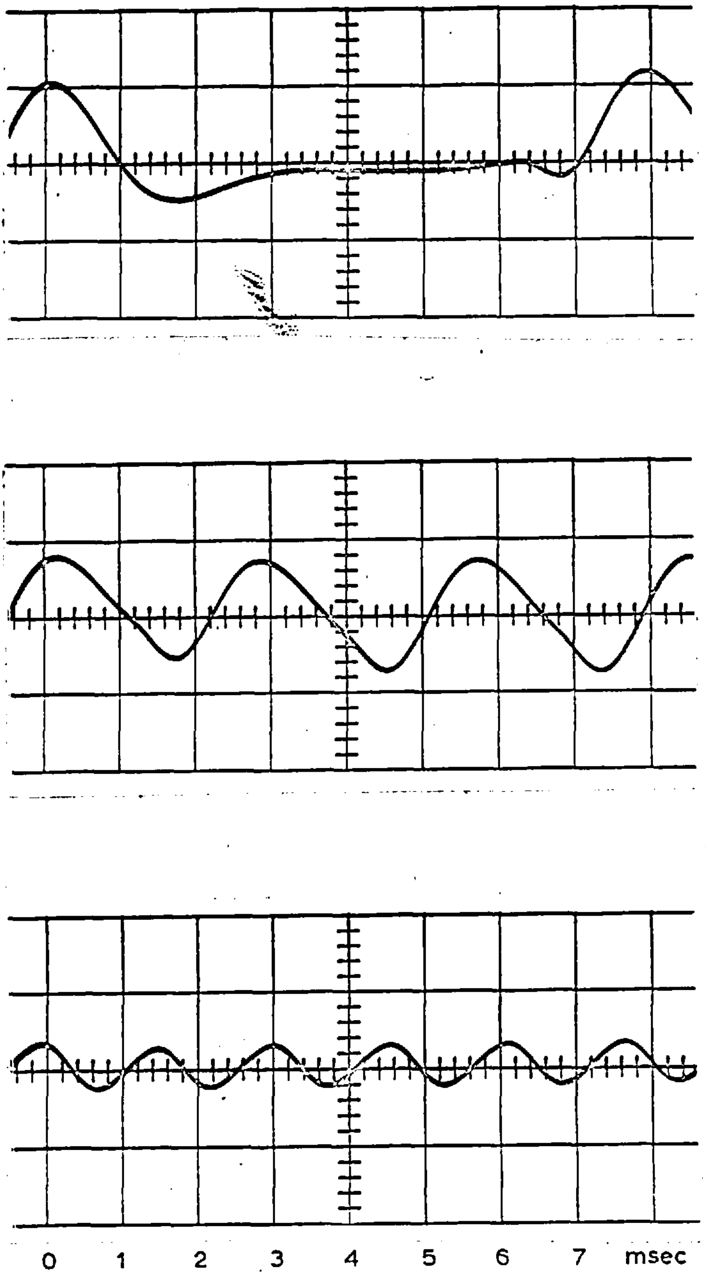


FIGURE 7.19

Effect of Repetition Period (T) on Basilar Membrane Response

Signal: Unfiltered, Repetitive Single Pulses



Condensations ↓
Waveform of
Basilar Membrane
Displacement at
300 cps Point

T = 8.0 msec
(0.5 volts/cm)

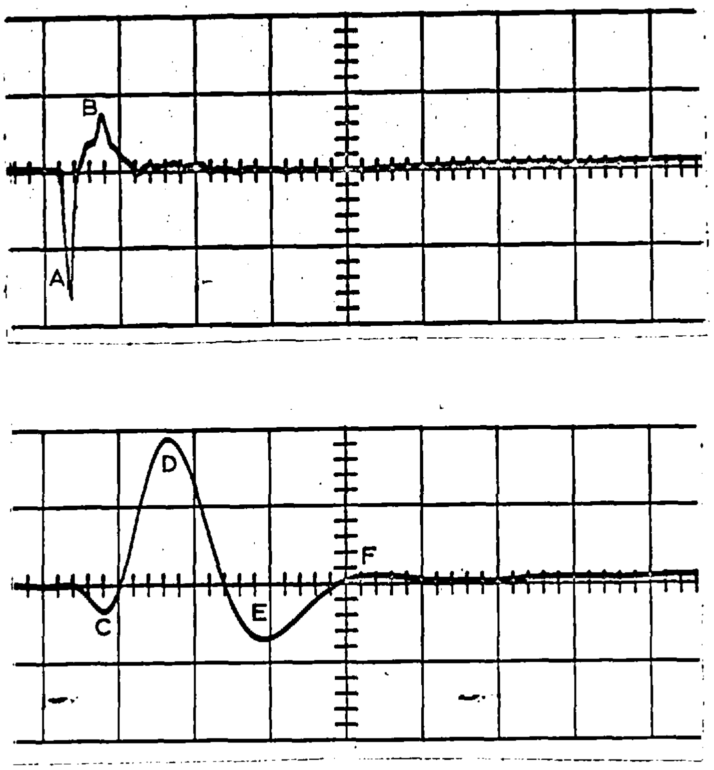
T = 3.0 msec
(0.5 volts/cm)

T = 1.5 msec
(0.5 volts/cm)

FIGURE 7.20

Basilar Membrane Response to Unfiltered Transient

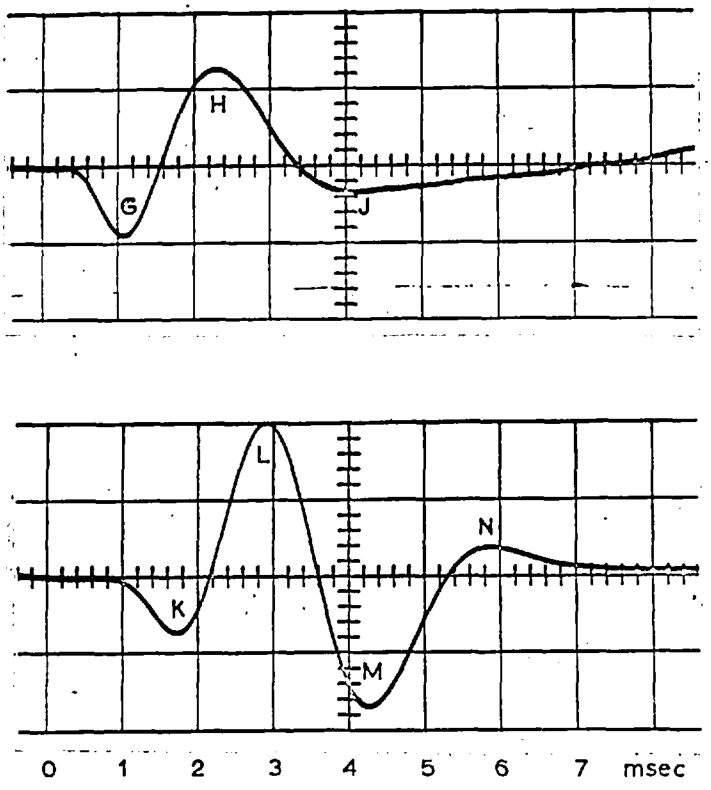
Condensations ↓



(a)
Waveform of
Acoustic Transient
(Unfiltered)

(b)
Waveform of
Basilar Membrane
Displacement at
400 cps Point

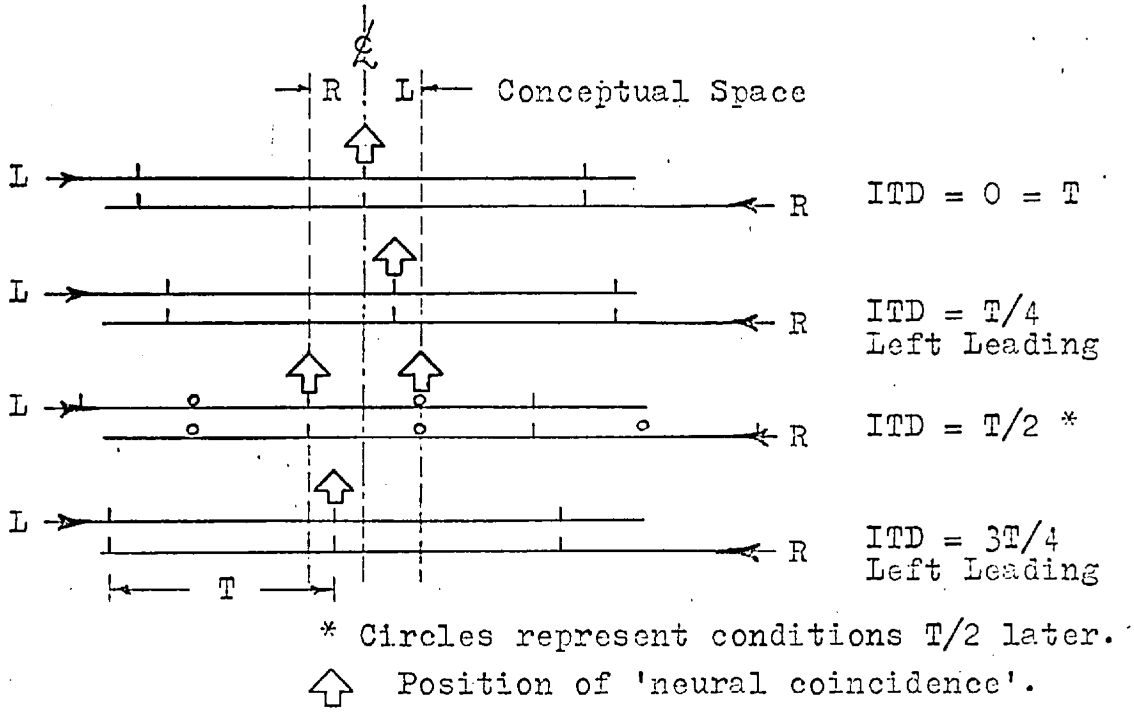
Basilar Membrane Response to Filtered Transient



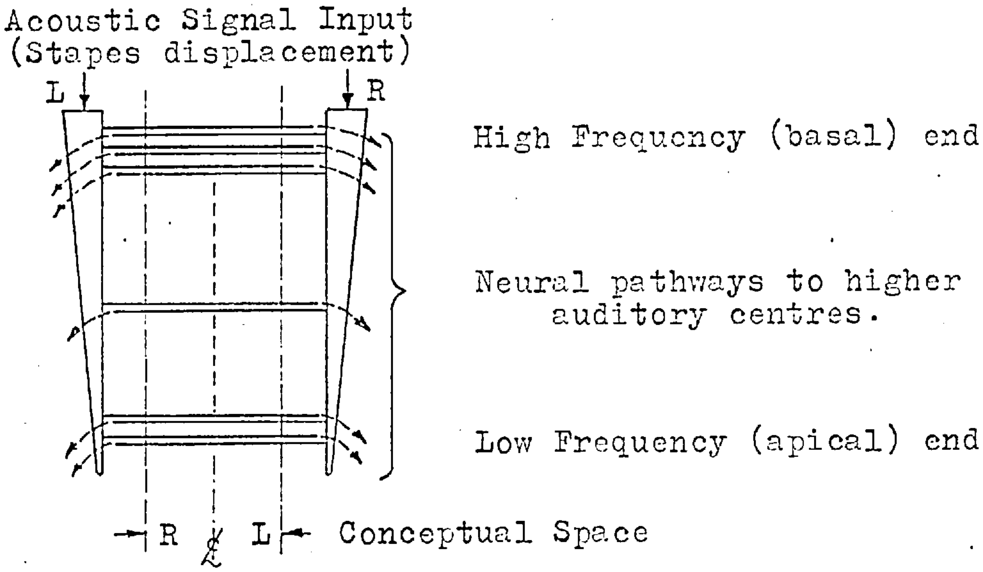
(c)
Waveform of
Acoustic Transient
(Filtered 20-400 cps)

(d)
Waveform of
Basilar Membrane
Displacement at
400 cps Point

FIGURE 7.21



(a) Illustration of interaural time to position transformation as effected by a simplified delay line/coincidence mechanism. Vertical lines represent bursts of neural activity associated with a repetitive feature of a binaural tone.



(b) Proposed working picture of the mechanism of cross-comparison.

Predictive image trajectories based on ITD's for centered images as listed in Table 7.2. Predictions were based on the cross-comparison of neural activity arising from corresponding regions of the cochleae, in this case the 400 cps points.

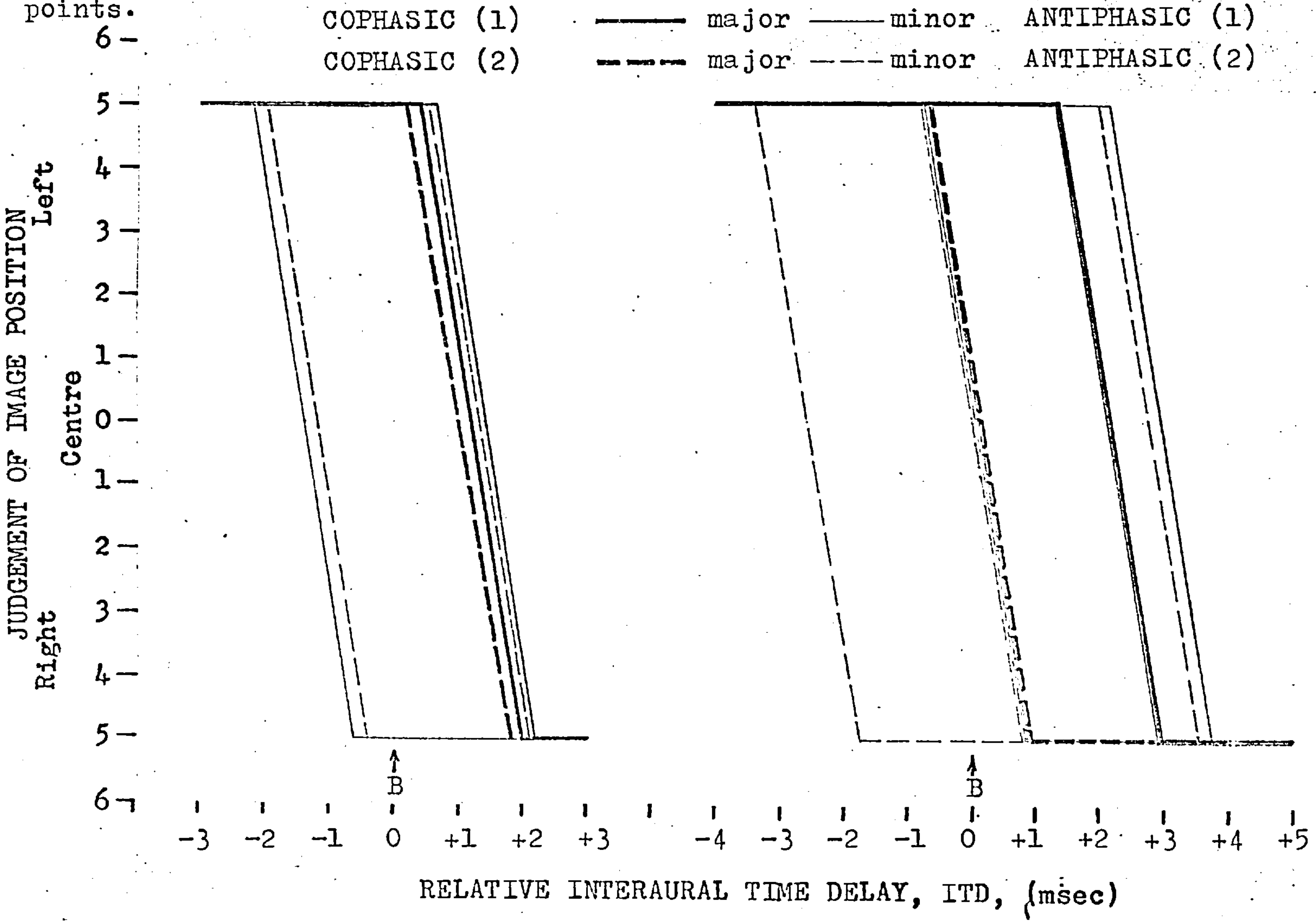
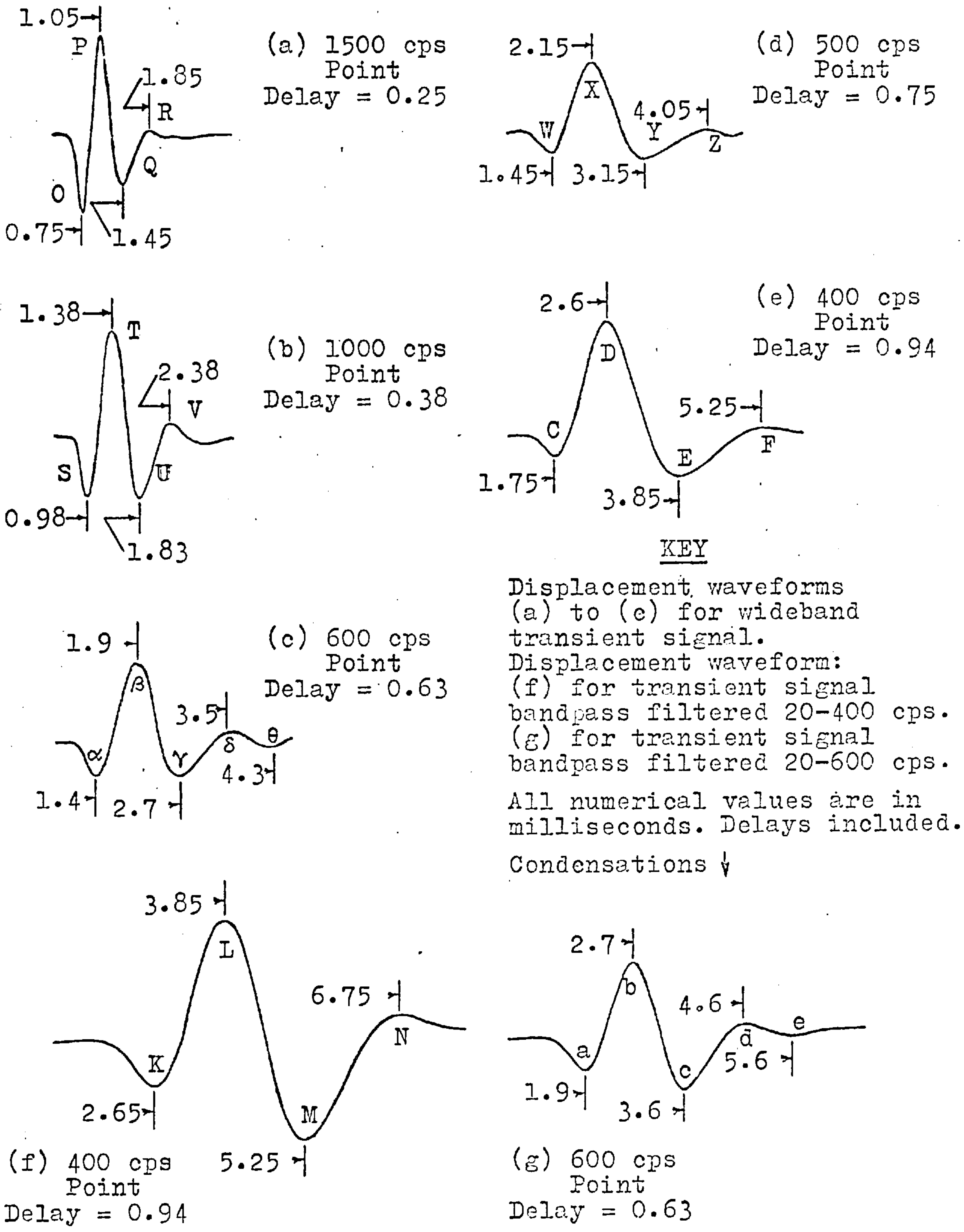


FIGURE 7.22



(a) 1500 cps
Point
Delay = 0.25

(d) 500 cps
Point
Delay = 0.75

(b) 1000 cps
Point
Delay = 0.38

(e) 400 cps
Point
Delay = 0.94

(c) 600 cps
Point
Delay = 0.63

KEY

Displacement waveforms (a) to (c) for wideband transient signal.
Displacement waveform:
(f) for transient signal bandpass filtered 20-400 cps.
(g) for transient signal bandpass filtered 20-600 cps.

All numerical values are in milliseconds. Delays included.

Condensations ↓

(f) 400 cps
Point
Delay = 0.94

(g) 600 cps
Point
Delay = 0.63

FIGURE 7.23

Bandwidth of Binaural Uncorrelated Masking Noise	SUBJECT: F.E.T.		SUBJECT: B.McA.S.		SUBJECT: D.G.P.	
	Cophasic	Antiphasic	Cophasic	Antiphasic	Cophasic	Antiphasic
600-15000 cps	+0.1	-0.9 +1.1	+0.15	-0.6 +0.9	-0.1	-1.1 +0.9
1000-15000 cps	Lower-Pitched Image +0.1	-0.9 +1.0	+0.12	-0.4 +0.7	-0.05	-0.5 +0.55
	Higher-Pitched Image +0.1	-0.55 +0.5				
2000-15000 cps	Lower-Pitched Image +0.05	-1.1 +1.1	+0.15	-0.4 +0.5	-0.05	-0.45 +0.5
	Higher-Pitched Image +0.05	-0.45 +0.5				
20-1500 cps	+0.1	+0.1	+0.15	+0.18	0.0	-0.02

TABLE 7.1 Table of ITD values (in msec) indicated by listeners as producing centered images. The experiments used unfiltered transient signals (T = 20 msec, A:B= ±1:1) which were bilaterally masked with filtered random noise. A positive delay indicates a lagging left-ear signal. Each result is the average of at least six separate centerings. One listener (F.E.T.) reported multiple-image centerings; the dominant image was that reported as having the 'higher-pitched' character.

CONDITION	MEASURED RESULTS		PREDICTIVE RESULTS	
	ITD values for Centered Images as Determined from Results of Image-Lateralization Experiments (msec)		Based on Cross-Comparison of Neural Activity Arising from Corresponding Regions: 400 cps Left, 400 cps Right.	Based on Cross-Comparison of Neural Activity Arising from Maximally Responding Regions: 1500 cps Left, 400 cps Right.
	F.E.T.	D.G.P.	ITD values for centered images. (msec)	ITD values for centered images. (msec)
Cophasic (1)	Left = Condensation		Right = Condensation	
Figs. 7.4 & 7.6	+1.1	+1.25	FL -1.4 DL +1.2 FN +1.4 DN +4.2	RL +2.0 LP +2.8 RN +4.9 PN +5.7
Cophasic (2)	Left = Rarefaction		Right = Rarefaction	
Figs. 7.5 & 7.7	+1.05	+1.15	EK -1.2 CK +1.0 EM +1.3 CM +3.5	KQ +1.2 KO +1.9 MQ +3.8 MO +4.5
Antiphasic (1)	Left = Rarefaction		Right = Condensation	
Figs. 7.4 & 7.6	+0.3 +1.8	+0.3 +2.4	EL 0.0 CL +2.1 EN +2.9 CN +5.0	LQ +2.4 LO +3.1 NQ +5.3 NO +6.0
Antiphasic (2)	Left = Condensation		Right = Rarefaction	
Figs. 7.5 & 7.7	+0.4 +1.5	+0.2 +2.1	FK -2.6 FM 0.0 DK +0.1 DM +2.7	KR +0.8 KP +1.6 MR +3.4 MP +4.2

TABLE 7.2 ITD values for centered images as measured from the results of image-lateralization experiments illustrated in Figs. 7.4, 7.5, 7.6, and 7.7 are compared with predictive values computed using the waveforms of Fig. 7.23. Letter designations associated with the predictive ITD's refer to the relevant peaks of basilar membrane displacements (Fig. 7.23). Positive ITD indicates left-ear signal lagging.

CONDITION	EXPERIMENTAL RESULTS			PREDICTIVE RESULTS			
Bandwidth of Masking Noise (cps)	ITD values for Centered Images as Determined in Image Centering Experiments (msec)			Based on Cross-Comparison of Neural Activity Arising from Corresponding Regions of Both Basilar Membranes		Based on Cross-Comparison of Neural Activity Arising from Maximally Responding Regions of Both Basilar Membranes	
				Points of Cross-Comparison Left - Right	Estimated ITD Values for Centered Images (msec)	Points of Cross-Comparison Left - Right	Estimated ITD Values for Centered Images (msec)
EXPERIMENT A	SIGNAL: T= 20 msec, A:B = ±5:±1, Bandpass Filtered: 20-400 cps, Left Channel only. MASKING NOISE: Right Channel Only, Filtered as Indicated.						
	F.E.T.	B.McA.S.	N.E.				
Cophasic (1)	Left = Condensation		Right = Condensation				
Unmasked	-0.90	-1.20	-1.00	400 - 400	DL FN FL DN -1.2 -1.4 +1.4 -4.2	400 - 1500	PN RN LP RL -5.7 -4.9 -2.8 -2.0
1000 - 15000	-0.90	-0.90	-1.00	400 - 400	DL FN FL DN -1.2 -1.4 +1.4 -4.2	400 - 1000	NT NV LT LV -5.4 -4.4 -2.5 -1.5
600 - 15000	-0.85	-0.90	-0.95	400 - 400	DL FN FL DN -1.2 -1.4 +1.4 -4.2	400 - 600	Nβ Nδ Lβ Lδ -4.9 -3.3 -1.9 -0.4
500 - 15000	-0.90	-0.90	-0.95	400 - 400	DL FN FL DN -1.2 -1.4 +1.4 -4.2	400 - 500	NX NZ LX LZ -4.6 -2.7 -1.7 +0.2
400 - 15000	-0.90	-0.90	-0.95	400 - 400	DL FN FL DN -1.2 -1.4 +1.4 -4.2	400 - 400	DL FN FL DN -1.2 -1.4 +1.4 -4.2
Cophasic (2)	Left = Rarefaction		Right = Rarefaction				
Unmasked	-0.90 -1.10	-1.30		400 - 400	CM EM CK EK -3.5 -1.3 -1.0 +1.2	400 - 1500	MO MQ KO KQ -4.5 -3.8 -1.9 -1.2
Antiphasic (1)	Left = Rarefaction		Right = Condensation				
Unmasked	-0.35 -2.00	-0.25 -1.80		400 - 400	CN EN CL EL -5.0 -2.9 -2.1 -0.0	400 - 1500	MP MR KP KR -4.2 -3.4 -1.6 -0.8
Antiphasic (2)	Left = Condensation		Right = Rarefaction				
Unmasked	-0.35 -2.10	-0.50 -2.10		400 - 400	DM DK FM FK -2.7 -0.1 0.0 +2.6	400 - 1500	NO NQ LO LQ -6.0 -5.3 -3.1 -2.4

Positive ITD indicates lagging left-ear signal.

TABLE 7.3

CONDITION	EXPERIMENTAL RESULTS		PREDICTIVE RESULTS			
Bandwidth of Masking Noise (cps)	ITD values for Centered Images as Determined in Image Centering Experiments (msec)		Based on Cross-Comparison of Neural Activity Arising from Corresponding Regions of Both Basilar Membranes		Based on Cross-Comparison of Neural Activity Arising from Maximally Responding Regions of Both Basilar Membranes	
			Points of Cross-Comparison Left - Right	Estimated ITD Values for Centered Images (msec)	Points of Cross-Comparison Left - Right	Estimated ITD Values for Centered Images (msec)
EXPERIMENT B	SIGNAL: T = 20 msec, A:B = ±5:±1, Bandpass Filtered: 20-600 cps, Left Channel Only. MASKING NOISE: None.					
	F.E.T.	B.McA.S.				
Antiphasic (1)	Left = Rarefaction	Right = Condensation				
Unmasked	-1.2 -0.1 +1.4	-1.4 -0.4	600 - 600	eδ -2.1 cβ -1.7 cδ -0.1 aβ 0.0 aδ +1.6	600 - 1500	eR -3.7 cP -2.5 cR -1.7 aP -0.9 aR 0.0
Antiphasic (2)	Left = Condensation	Right = Rarefaction				
Unmasked	-1.8 -1.1 -0.1 +1.5	-1.7 -0.5 +1.5	600 - 600	αd -3.2 γd -1.9 αb -1.3 θd -0.3 γb 0.0 θb +1.6	600 - 1500	dO -3.8 dQ -3.1 bO -1.9 bQ -1.2
EXPERIMENT C	SIGNAL: T = 20 msec, A:B = ±1:±1, Unfiltered. MASKING NOISE: Right Channel Only, Filtered (Bandpass) as Indicated.					
Cophasic (1)	Left = Condensation	Right = Condensation				
Unmasked	-0.0	B.McA.S.	1500 - 1500	RP -0.8 PP 0.0 RR 0.0 PR +0.8	1500 - 1500	RP -0.8 PP 0.0 RR 0.0 PR +0.8
1000 - 15000	+0.2 +1.0 (+0.1)	B.McA.S. (Table 7.1)	1000 - 1000	VT -1.0 TT 0.0 VV 0.0 TV +1.0	1500 - 1000	RT -0.5 PT +0.3 RV +0.5 PV +1.3
600 - 15000	+0.2 +1.4 (+0.2)	B.McA.S. (Table 7.1)	600 - 600	δβ -1.6 ββ 0.0 δδ 0.0 βδ +1.6	1500 - 600	Rβ +0.1 Pβ +0.9 Rδ +1.7 Pδ +2.5
400 - 15000	+0.1	B.McA.S.	400 - 400	FD -2.7 DD 0.0 FF 0.0 DF +2.7	1500 - 400	RD +0.8 PD +1.6 RF +3.4 PF +4.2

Positive ITD indicates lagging left-ear signal.

TABLE 7.4

CONDITION	EXPERIMENTAL RESULTS						PREDICTIVE RESULTS									
Bandwidth of Masking Noise (cps)	ITD values for Centered Images as Determined in Image Centering Experiments (msec)						Based on Cross-Comparison of Neural Activity Arising from Corresponding Regions of Both Basilar Membranes					Based on Cross-Comparison of Neural Activity Arising from Maximally Responding Regions of Both Basilar Membranes				
							Points of Cross-Comparison Left - Right	Estimated ITD Values for Centered Images (msec)				Points of Cross-Comparison Left - Right	Estimated ITD Values for Centered Images (msec)			
EXPERIMENT C Continued	SIGNAL: T = 20 msec, A:B = ±1:±1, Unfiltered. MASKING NOISE: Right Channel Only, Filtered (Bandpass) as Indicated.															
Antiphasic (1)	Left = Rarefaction				Right = Condensation											
	Low	Med.	High	Med.	Low*											
Unmasked	-1.1	-0.5	-0.2	+0.2	+1.2	F.E.T.	1500 - 1500	PQ -0.4	PO +0.3	QR +0.4	OR +1.1	1500 - 1500	PQ -0.4	PO +0.3	QR +0.4	OR +1.1
				+0.2		B.McA.S.	400 - 400	DE -1.3	DC +0.9	EF +1.4	CF +3.5					
1000 - 15000	-1.1 (-0.9)	-0.6 (-0.6)		+0.3 (+0.5)	+1.0 (+1.0)	F.E.T. (Table 7.1)	1000 - 1000	TU -0.5	TS +0.4	UV +0.6	SV +1.4	1500 - 1000	TQ -0.1	OT +0.6	QV +0.9	OV +1.6
		-0.4 (-0.4)		+0.7 (+0.7)	+1.3	B.McA.S. (Table 7.1)	400 - 400	DE -1.3	DC +0.9	EF +1.4	CF +3.5					
600 - 15000	-1.1 (-0.9)	-0.5		+0.5	+1.0 (+1.1)	F.E.T. (Table 7.1)	600 - 600	βγ -0.8	βα +0.5	γδ +0.8	αδ +2.1	1500 - 600	Qβ +0.5	Oβ +1.2	Qδ +2.1	Oδ +2.8
		-0.6 (-0.6)		+0.8 (+0.9)		B.McA.S. (Table 7.1)	400 - 400	DE -1.3	DC +0.9	EF +1.4	CF +3.5					
400 - 15000	-1.0				+1.0	F.E.T.	400 - 400	DE -1.3	DC +0.9	EF +1.4	CF +3.5	1500 - 400	QD +1.2	OD +1.9	QF +3.8	OF +4.5
		-0.6		+0.8		B.McA.S.										
Antiphasic (2)	Left = Condensation				Right = Rarefaction											
Unmasked	-1.1	-0.5	0.0	+0.5	+1.2	F.E.T.	1500 - 1500	OR -1.1	RQ -0.4	PO -0.3	PQ +0.4	1500 - 1500	OR -1.1	RQ -0.4	PO -0.3	PQ +0.4
				+0.2		B.McA.S.(1)	400 - 400	CF -3.5	EF -1.4	DC -0.9	DE +1.3					
	-0.9	-0.4	+0.1	+0.4		B.McA.S.(2)										
600 - 15000	-1.1	-0.5		+0.6	+1.4	F.E.T.	600 - 600	αδ -2.1	γδ -0.8	βα -0.5	βγ +0.8	1500 - 600	Rα -0.5	Pα +0.4	Rγ +0.9	Pγ +1.7
		-0.5		+0.6		B.McA.S.	400 - 400	CF -3.5	EF -1.4	DC -0.9	DE +1.3					

*Apparent pitch of the centered impulsive image.
Reported only by the listener F.E.T.

Positive ITD indicates lagging left-ear signal.

TABLE 7.5

CHAPTER 8

CONCLUSIONS

The important findings of the present study may be summarized as follows:

(1) it is entirely feasible to require subjects to report judgements of the degree of apparent lateral displacement of the binaurally created and perceived acoustic images which arise in earphone listening. Such positional judgements were found to be reproducible and generally of the same form among listeners.

(2) when signals giving rise to perceptions of acoustic images were presented and judged in conjunction with signals producing visual images of a compatible nature, a number of cross-modal effects could be demonstrated. Certain of the effects were of a transitory nature and could be described in terms of dominance of the visual image positional information; however, an effect of a semi-permanent nature was observed in which the relationship between aural signal ITD and judged image position was altered. The alteration was not always observed, and when it was observed it was not always of a form which was in a simple relation to the bimodal experience. It was clear that prior to further experiments of this type, more knowledge must be made available as to the characteristics of the processes underlying acoustic image phenomena, or, in other words, of the basic combination of binaural auditory signals.

(3) the patterns of judged image position versus ITD (judged image trajectories) that resulted from experiments conducted on the basis of (1) were found to be of a convenient form for assisting interpretation of a number of complex acoustic image phenomena.

Firstly, basic aspects of Wallach's precedence effect were found to be well accounted for in terms of positional judgements in situations with multiple coexisting images. Secondly, in examining monaural temporal masking effects, it was established that the isolated use of an image-centering technique could lead to results which may be easily misinterpreted. Moreover, it was found that the degree to which monaural masking effects are evident in judged image trajectories (where appropriate) is not only a function of acoustic signal configuration, but also of the image which is being judged, and of the underlying cochlear mechanism.

(4) a further finding of the present study was that in experiments using complex acoustic stimuli (e.g. transients) situations commonly arose in which several images could be simultaneously identified. It was found that where listeners were required simply to lateralize an undifferentiated image, these coexisting images could, at different times, contribute differently to the lateralization judgements. Experiments in which listeners were required to classify and independently judge the various images were useful in resolving such results. Judged trajectories of these images were such that in many cases it was possible objectively to assess the subjective image identification. It is argued that substantial clarification of judgements in these involved situations has been achieved in the following terms.

The simultaneously perceived images were found to fit into two categories: multiple images, i.e., those which had distinctly different pitches or timbres, such that listeners could reliably identify them on different occasions, and split images, i.e., those which were reported as having closely similar or identical identifying characteristics, but which were separated in judged position.

Positional judgements pertaining to split images were observed to fall into groups clearly related to either or both of the split image elements, or to some intermediate position suggesting a balance of attention between the two similar sounding, but spatially distinct images.

In experiments using binaural repetitive acoustic transients the dominant image was commonly reported as having strongly impulsive character; however, in certain conditions, multiple images of tonal character could also be identified. These tonal images were clearly related to the lower harmonics of the pulse repetition frequency, and were concluded to be the result of a spectral analysis or selection of the binaural repetitive transient signal. Many listeners were able to report only a single (dominant) impulsive image in experiments with wideband transients; however, a few listeners could identify two impulsive images, and one listener regularly reported three images of impulsive character. Judged trajectories of these images indicated that they were associated with temporal features of the acoustic signal as represented in basilar membrane displacements.

(5) It was found possible to ascertain the probable cochlear origins of the multiple impulsive images; this study was based on observations of detailed responses of the basilar membrane (as indicated by a model) to measured acoustic transients. The multiple impulsive images were shown by preliminary masking experiments to be associated with neural activity arising from specific regions of the cochlea.

It was assumed that in the cochlea, the process of mechanical-to-neural transduction was of a simple (and reasonable) kind, viz., that neural firings occur only on upward displacements of the basilar membrane (due to an outward movement of the stapes as would occur for a pressure rarefaction). Further, it was assumed that there is an equivalent time-to-space transformation performed on the neural concomitants of the binaural acoustic signals, such that (in this case) ITD is represented by a factor which parallels image displacement in the conceptual space in which images are perceived.

On the basis of these two assumptions (for which some neurophysiological evidence has been brought forward) and measurements of the time course of inferred basilar membrane displacements, a working picture of the mechanism of binaural interaction was developed, and predictive results tested against the results of specially designed image-lateralization and image-centering experiments. It was found that a working picture based on multiple cross-comparisons (i.e. elemental time-to-space transformations) of neural impulses arising from corresponding narrow regions of the cochleae provided consistent and reasonably precise descriptions and predictions of experimental results for a variety of signals.

The region of the cochlea contributing to the important cross-comparison (i.e. dominant image) was found to be determined by two factors: (1) the nature of the acoustic transient signal (i.e., bandwidth, repetition rate) which establishes the overall pattern of basilar membrane displacement and (2) the masking conditions which establish the regions of the cochlea from which significant coherent neural activity can arise. In the absence of masking noise, the dominant cross-comparison was deduced as being effected between corresponding regions of the cochleae which were associated, in either or both of the cochleae, with a maximum or peak of cochlear mechanical activity. In the presence of masking noise the cross-comparison giving rise to the dominant image, was apparently associated with corresponding unmasked regions of the cochleae either or both of which were producing maximal coherent neural activity (again deduced on the basis of inferred basilar membrane displacement amplitude).

It was concluded that the remaining two impulsive images were associated with the basal (high-pitched image) and extreme apical (low-pitched image) regions of the cochlea, but the factors underlying their independent identification were less clear.

APPENDIX A1

APPARATUS FOR GENERATING BINAURAL PULSE SIGNAL

Fig. A1.1 illustrates the scheme used for generating the brief electrical pulses which, when applied to binaural earphones, provided the experimental acoustic transient signals. The output from a 100 cps 'clock' multivibrator (waveform 1) provided synchronizing pulses to initiate the fixed delay (waveform 5) generator and was used also to gate a linear sweep generator (waveform 2). The trailing edge of the fixed-delay voltage was sensed and used to trigger the signal pulse generator which provided drive for one earphone (Output 1). The control voltage from the magnetic tape recorder (waveform 4) and the sweep voltage (waveform 2) were applied to a voltage comparator which sensed the negative-going point of equality of the two signals and generated a brief transient (point 3) at that instant. Changes in the control voltage altered the time at which voltage coincidence occurred and thereby the relative time at which the Output 2 signal pulses were generated. Suitable adjustment of the magnitude of the fixed delay and of the quiescent control voltage enabled the temporal spacing of Output 1 and Output 2 signal pulses (i.e. ITD) to be altered linearly with variations in the ITD control voltage.

The apparatus was designed using standard circuit techniques and was fully transistorized.

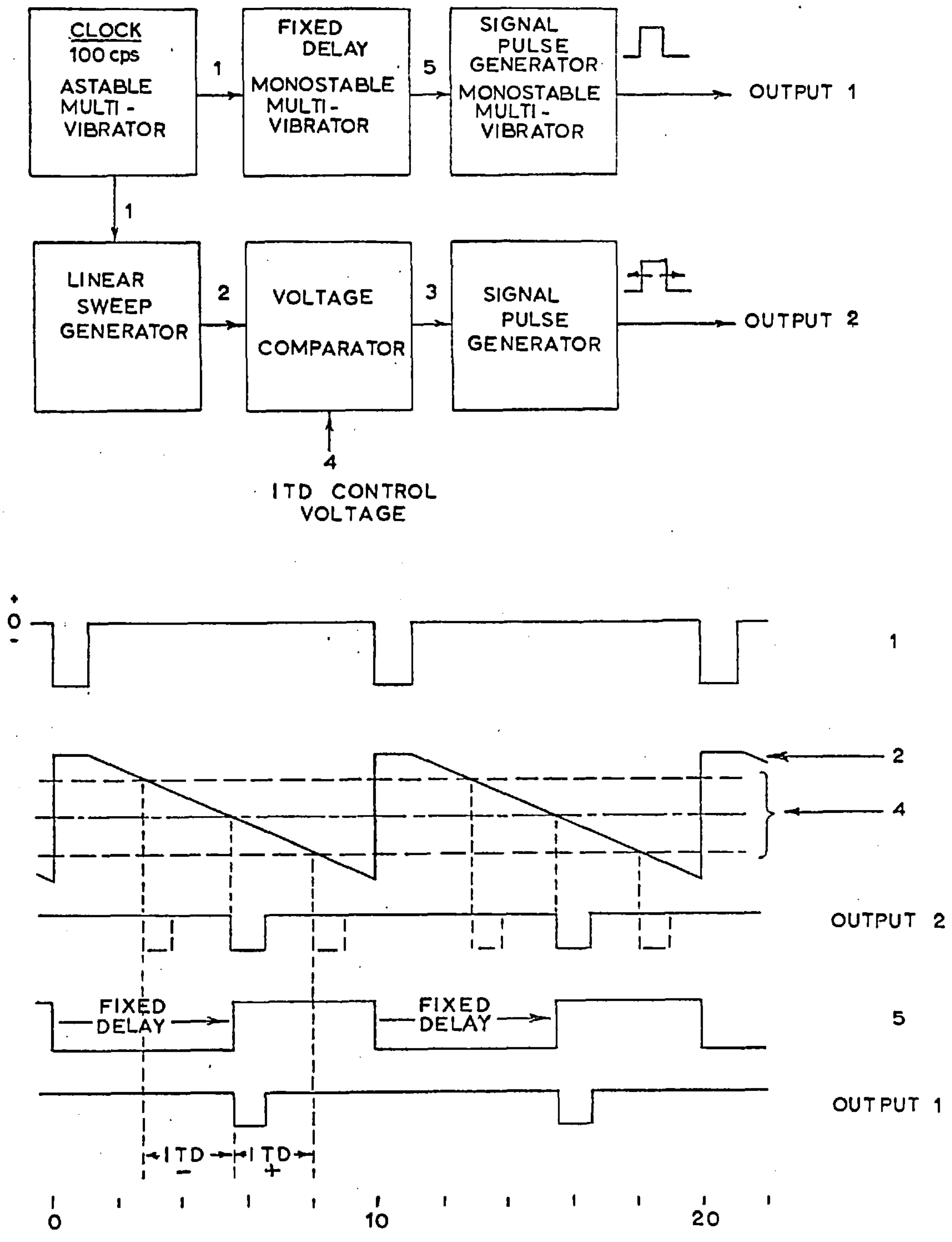


FIGURE A1.1 The scheme employed for generating binaural pulse signals with voltage-controlled ITD.

APPENDIX A2

EQUIPMENT FOR CONTROL OF AUTOMATED EXPERIMENTS

A2.1 Magnetic Tape Recording Apparatus

The equipment for the recording and playback of variable-voltage signals for use in controlling aural and visual signal parameters was constructed around a Leevers-Rich type E101R tape transport mechanism fitted with three Epsylon Industries heads arranged for parallel four-track erase, record and playback; see Fig. A2.1.

ERASURE: A 100 kilocycle oscillator provided drive to the erase head for the erasure of signals on any or all of the four tracks.

RECORDING: the basis of the recording scheme is the translation of the variable input voltage to the variable width of 1000 cps pulses. This was achieved in the Pulse Width Modulator (basically an emitter-coupled transistor monostable multivibrator with controlled base voltage on one transistor) which provided output pulses of width which could be altered linearly with input voltage over a range of about 0.4 msec. The Record Amplifier applied a differentiated version of this signal to the record head of the tape transport.

PLAYBACK: the recorded signal was tailored so that the signal which appeared at the playback head was of a form which could be accurately decoded and which was not susceptible to tape 'drop-outs' (i.e. sudden variations in recorded signal strength due to tape imperfections or particles of dust on the tape). This waveform is shown at the top of Fig. A2.1. The important feature of this waveform is the rapid transition between successive positive - and negative - going peaks. After amplification and clipping, differentials of these rapid voltage changes were used to trigger a bistable multivibrator the output of which was a close reconstitution of the original pulse-width modulator output. Demodulation (i.e. low-pass filtering) completed the sequence of operations and provided an output signal which was proportionately similar to the input signal originally recorded.

A2.2 The Recording of an Experiment Control Programme

In the experiments reported in this study the proportional changes of aural and visual signal parameters were identical (only fixed biases were introduced). Therefore, it was necessary to record only one signal-parameter control voltage and the two-level voltage for controlling the application and removal of aural and/or visual signals.

To record an experiment-control programme with randomized parameter changes and 4-second signal presentations, the procedure was as follows: a low-frequency square-wave generator (8 second period, 1:1 mark-to-space ratio) supplied a two-level voltage to one recording channel). During the 4 second intervals in which the experimental signals would be removed, the 11-level parameter-control voltage (applied to a second recording channel) was adjusted according to a predetermined randomized sequence of 110 values.

A preparatory sequence of nominal centre, + maximum and - maximum, each repeated five times, was recorded immediately prior to the experimental sequence described above. As well as being used for purposes of introducing subjects to the range of signal variation, this sequence was useful for purposes of equipment calibration before each experiment.

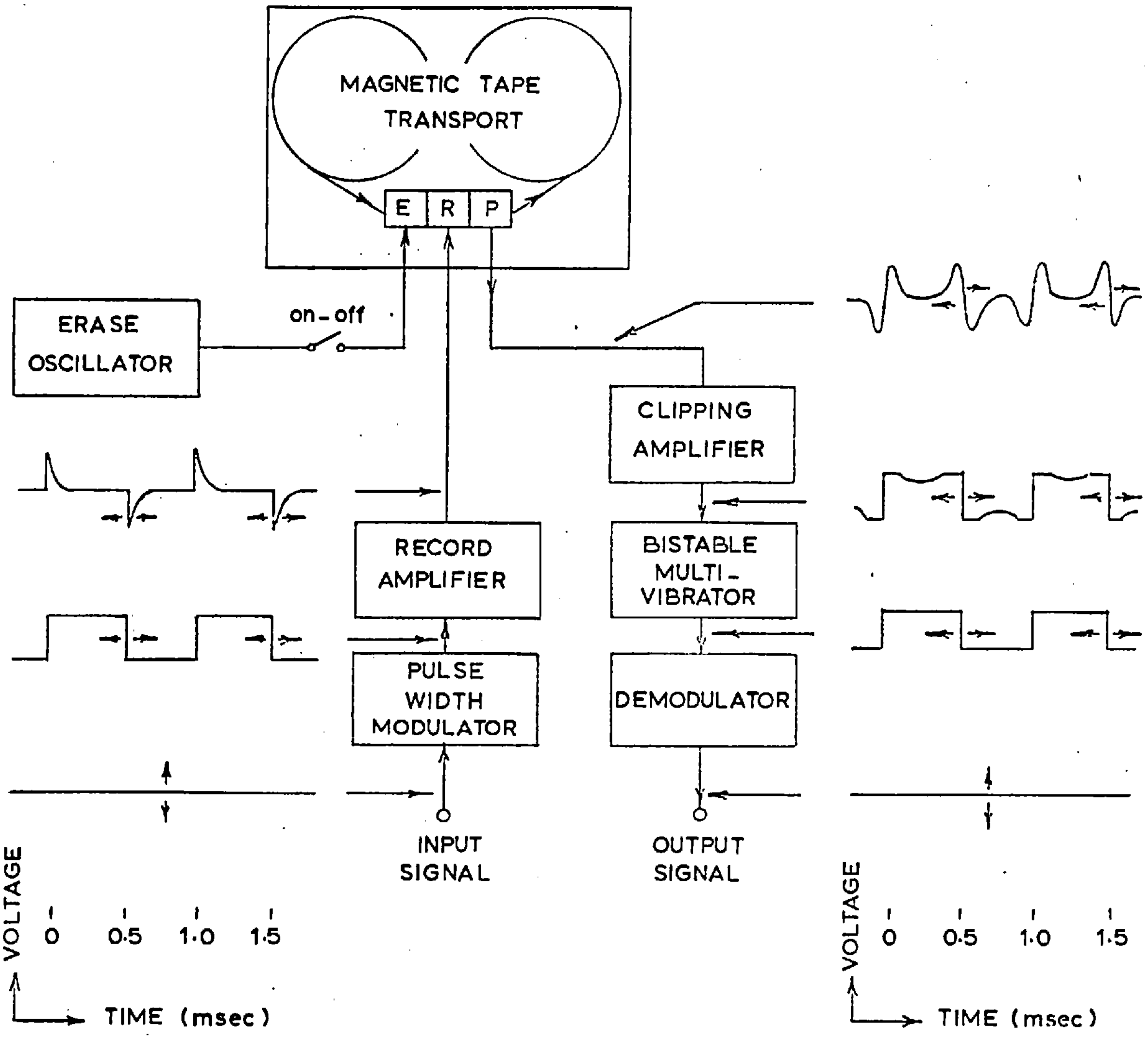


FIGURE A2.1 The scheme employed for the recording and playback of signals for use in controlling aural and visual signal parameters. Waveforms depict conditions at various points in the record and playback chains. Equipment for one of the four channels has been illustrated.

APPENDIX A3

EQUIPMENT FOR RECORDING EXPERIMENTAL DATA

As illustrated in Fig. 2.1, the experimental data was printed out in two forms: numerical (punched tape) and graphical (X-Y plotter). The "Sample Control," which sensed the removal of the experimental signals, triggered the digital voltmeter (Solartron LM.902) which converted the analog voltage from the response indicator into digital form. The parallel decade coded output from the digital voltmeter was applied to an encoder (Solartron LP.983) which serialized and encoded the signal and provided the necessary drive for the paper tape punch (Creed Model 25).

To enable the data on the punched tapes to be converted to convenient numerical form in a teleprinter it was necessary that occasional 'line-feed' (LF) and 'carriage -return' (CR) characters be introduced. For this purpose an additional encoder was constructed which, after each number (word) punched by the Solartron encoder, added either two 'space' characters or LF and CR characters. The encoder was arranged such that it normally added the two 'space' characters, but at selected intervals (after every 1, 2 or 4 words) it would add LF + CR. Thus the typewritten version of the punched tape would have the numerical data conveniently arranged on the page in either 1, 2 or 4 columns.

Based on knowledge of the recorded signal-parameter sequence, masks were made which were placed over the typed copy to show the subject's judgements, in order, for each of the 11 signal-parameter values. The graphical data recording was achieved using a Houston Instrument Corporation EHR-93, X-Y plotting device, with modified pen-drop mechanism.

LIST OF REFERENCES

ALLANSON, J.T. and I.C. WHITFIELD (1956) The cochlear nucleus and its relation to theories of hearing; in C. Cherry (ed.) Information Theory, Third London Symposium, 1955. Butterworths Scientific Publications, London.

von BEKESY, G. (1951) Microphonics produced by touching the cochlear partition with a vibrating electrode. J. Acoust. Soc.Amer., 23, 29-35.

von BEKESY, G. (1930) Zur Therie des Horens, Physikal Zsch., 31, 857-858.

von BEKESY, G. (1953) Shearing microphonics produced by vibrations near the inner and outer hair cells. J. Acoust. Soc. Amer., 25, 786-790.

von BEKESY, G. (1960) Experiments in Hearing. Edited by E.G. Wever, McGraw-Hill Company, Inc., New York.

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

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

DAVID, E.E. Jr., N. GUTTMAN,W.A. van BERGEIJK (1958) On the mechanism of binaural fusion. J. Acoust. Soc. Amer., 30, 801-802.

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

DAVIS, H., I. TASAKI and R. GOLDSTEIN (1952) Peripheral origin of activity with reference to the ear. Cold Spring Harbour Symposia Quant. Biol., 17, 143-154.

DAVIS, H. (1957) Biophysics and physiology of the inner ear. Physiological Reviews, 37, 1-49.

DAVIS, H. (1958) A mechano-electrical theory of cochlear action. Ann. Oto. Rhin. Laryn., 67, 789-801.

DAVIS, H. (1959) Excitation of auditory receptors. Handbook of Neurophysiology, Vol. 1, American Physiological Society, 1959.

DAVIS, H. (1961) Some principles of sensory receptor action. Physiolog. Rev., 41, 391-416.

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

DEATHERAGE, B.H., D.H. ELDREDGE and H. DAVIS (1959) Latency of action potentials in the cochlea of the guinea pig. J. Acoust. Soc. Amer., 31, 479-486.

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

FLANAGAN, J.L. (1960) Models for approximating basilar membrane displacement. Bell System Technical Journal, 39, 1163-1191.

FLANAGAN, J.L. (1962) Models for approximating basilar membrane displacement - Part 2. Bell System Technical Journal, 41, 959-1009.

FLANAGAN, J.L., E.E. DAVID, Jr. and B.J. WATSON (1964) Binaural lateralization of cophasic and antiphasic clicks. J. Acoust. Soc. Amer., 36, 2184-2193.

FRANSSEN, N.V. (1964) Stereophony. Philips Technical Library, Cleaver-Hume Press, Ltd., London.

GALAMBOS, R. and H. DAVIS (1943) Response of single auditory-nerve fibers to acoustic stimulation. J. Neurophysiol., 6, 39-57.

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

GUTTMAN, N. and J.L. FLANAGAN (1964) Pitch of high-pass-filtered pulse trains. J. Acoust. Soc. Amer., 36, 757-765.

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

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

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

JACKSON, C.V. (1953) Visual factors in auditory localization. Quart. Jour. Exp. Psychol., 52-65.

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

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

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

LICKLIDER, J.C.R. (1951) Basic correlates of the auditory stimulus; in S.S. STEVENS (ed.) Handbook of Experimental Psychology, 985-1039. Wiley & Sons, New York.

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.

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

NORDMARK, J. (1963) Some analogies between pitch and lateralization phenomena. J. Acoust. Soc. Amer., 35, 1544-1547.

PLOMP, R. (1964) The ear as a frequency analyzer. J. Acoust. Soc. Amer., 36, 1628-1636.

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

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

SAYERS, B.Mc.A. (1964) Acoustic-image lateralization judgements with binaural tones. J. Acoust. Soc. Amer., 36, 923-926.

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.

SMALL, A.M. Jr. and R.A. CAMPBELL (1961) Masking of pulsed tones by bands of noise. J. Acoust. Soc. Amer., 33, 1570-1576.

SUMBY, W.H. and I.POLLACK (1954) Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Amer., 26, 212-215.

STEVENS, S.S. (1951) ed. Handbook of experimental psychology. Wiley and Sons, New York.

TASAKI, I., H. DAVIS and J-P. LEGOUIX (1952) Space-time pattern of cochlear microphonics (guinea pig) as recorded by differential electrodes. J. Acoust. Soc. Amer., 24, 502-519.

TASAKI, I. (1954) Nerve impulses in individual auditory nerve fibers of guinea pig. J. Neurophysiology, 17, 97-122.

TEAS, D.C. (1962) Lateralization of acoustic transients. J. Acoust. Soc. Amer., 34, 1460-1465.

TEAS, D.C., D.H. ELDREDGE, and H. DAVIS (1962) Cochlear responses to acoustic transients: an interpretation of whole-nerve action potentials. J. Acoust. Soc. Amer., 34, 1438-1459.

TONNDORF, J. (1962) Time/frequency analysis along the partition of cochlear models: a modified place concept. J. Acoust. Soc. Amer., 34, 1337-1350.

WALLACH, H., E.B. NEWMAN and M.R. ROSENZWEIG (1949) The precedence effect in sound localization. Amer. J. Psychol., 62, 315-336.

WEISS, T.F. (1964) A model for firing patterns of auditory nerve fibers. Technical Report 418, Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, Mass.

WHITWORTH, R. and L.A. JEFFRESS (1961) Time vs. intensity in the localization of tones. J. Acoust. Soc. Amer., 33, 925-929.

WOODWORTH, R.S. and H. SCHLOSBERG (1955) Experimental psychology (Revised edition). Methuen and Co. Ltd., London.

ZWICKER, E., G. FLOTTORP and S.S. STEVENS (1957) Critical band width in loudness summation. J. Acoust. Soc. Amer., 29, 548-557.