




114
179
THS



This is to certify that the
thesis entitled
**Perception of Frequency Modulation
width at Low Modulation Frequencies**

presented by
Mark Allen Klein

has been accepted towards fulfillment
of the requirements for
M.A. degree in Psychology


Major professor

Date Winter 1980

PERCEPTION OF FREQUENCY
MODULATION WIDTH AT
LOW MODULATION
FREQUENCIES

By

Mark Allen Klein

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1980

ABSTRACT

PERCEPTION OF FREQUENCY MODULATION WIDTH AT LOW MODULATION FREQUENCIES

By

Mark Allen Klein

Previous studies have shown that the perceived width of frequency modulation depends not only on the frequency excursion, $\pm f$, but also on the modulation rate and waveform. The results of this earlier work failed to support a single model of modulation width perception. In this study a series of experiments was carried out based on a complex modulation waveform having only two spectral components. Parametric manipulations involved the overall frequency excursion, the phase relationship between the two components and the modulation rate. Both modulation width perception and modulation waveform discrimination showed dependence on these parameters. The class of models for width perception involving simple detection of spectral components of the modulation waveform, ignoring phase relationships, was rejected. Future models must more closely examine the temporal microstructure of the modulation waveform.

ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. James Zacks, Dr. Gordon Wood, and Dr. David Wessel for their help in serving as members of my thesis committee. I am also very grateful to Dr. William Hartmann, my major advisor, whose guidance, assistance, and inspiration made this work possible. My special thanks to my wife, Diane, for her help in preparing this manuscript and understanding support throughout the entire duration of this project.

TABLE OF CONTENTS

LIST OF TABLESiv
LIST OF FIGURES.	v
Views of Perception.	4
Models	7
Experiment I11
Method13
Results-Discussion18
Experiment II.35
Method36
Results-Discussion37
Experiment III38
Method39
Results40
Discussion40
Experiment IV.44
Method49
Results-Discussion49
General Discussion and Conclusions53
LIST OF REFERENCES56

LIST OF TABLES

Table 1.	Predicted and observed R_e -values.	9
Table 2.	95% confidence intervals for differences between pooled data for $\phi = \pi/4, \pi/2, \& 3\pi/4$ versus pooled data for $\phi = 5\pi/4, 3\pi/2, \& 7\pi/4..$31
Table 3.	95% confidence intervals for differences between Q_e -values across levels of $\Delta f_g..$32
Table 4.	FM detection levels in Hertz for 4 and 12 Hz modulation rates.42
Table 5.	95% confidence intervals for differences in Q_e -values for $\phi < \pi$ versus $\phi > \pi..$51

LIST OF FIGURES

Figure 1.	Example waveforms.	5
Figure 2.	Complex waveforms.12
Figure 3.	Trial timing17
Figure 4.	R_e -values. Group data Experiment I. For this and all following figures \square represents $\Delta f = 2.5$ Hz Δ represents $\Delta f_s = 5$ Hz and \circ represents $\Delta f_s = 10$ Hz .19	.19
Figure 5.	R_e -values. Subject B Experiment I20
Figure 6.	R_e -values. Subject M Experiment I21
Figure 7.	R_e -values. Subject W Experiment I22
Figure 8.	R_e -values. Auxillary subjects Experiment I.23
Figure 9.	Q_e -values. Group data Experiment I and II25
Figure 10.	Q_e -values. Subject B Experiment I and II26
Figure 11.	Q_e -values. Subject M Experiment I and II27
Figure 12.	Q_e -values. Subject W Experiment I and II28
Figure 13.	Q_e -values. Auxillary subjects Experiment I.29
Figure 14.	Amplitude of 12 Hz component relative to thresholds.41
Figure 15.	Exponential windows superimposed on modulation waveforms .47	
Figure 16.	Q_e -values. Experiment IV data.50

Vibrato is a common musical ornament used in singing and in playing most musical instruments. The basic mechanism of vibrato is a slow, periodic modulation of the frequency being produced about some central frequency. Vibrato fulfills various roles for the performer. Kock (1936) found vibrato effective for disguising incorrect intonation in the component tones of chords. Sundberg (1978), however, found that vibrato does not disguise incorrect intonation for single tones. An earlier Sundberg paper (1977) found that vibrato does conceal deviations in vowel qualities typical in singing. The artistic role of vibrato includes, for example, variation of emotional expression for single notes or phrases. The performer may vary the depth and/or rate of vibrato during a single note or phrase to achieve a desired effect. When specifically notated, vibrato may have any other meaning or effect depending on the abilities of the composer/arranger and the performer.

Two entire volumes (C.E. Seashore, 1932, 1935) were devoted to discovering the depth and rate for the most pleasing and beautiful vibrato. Techniques for production of this ideal vibrato were analyzed and methods for teaching those techniques to other musicians were described. Most of those studies, however, offered no information on how vibrato is perceived; how frequency modulation is processed by the auditory system.

As with phenomena in other complex systems, vibrato in musical instruments and the human voice is much more than a single, simple effect i.e., frequency modulation (FM). There are unavoidable changes in the amplitude and/or timbre of the tone, too. These effects may be due to the technique for production of the vibrato or the acoustical properties of the instrument or vocal tract. Despite these other changes,

the major effect perceived is the modulation of the pitch. Traditionally vibrato is frequency modulation. In this study, vibrato will be represented by the process of frequency modulation of a sinewave carrier signal.

Vibrato and frequency modulation are not isolated musical phenomena. They are part of the more general problem of the perception of pitch and pitch changes over time. Zwicker (1973) has argued that the perception of frequency modulation may encompass all the elements of time-varying pitch perception. The auditory processes involved in time-varying pitch perception may also extend to the perception of other musical ornaments, such as trills (e.g., Dowling, 1973), or even the basic phenomena of melody perception (e.g., Deutsch, 1972). In speech perception, pitch levels and especially transitions between pitch levels are important for carrying information and characterizing phonemes. The typical model system for studying these effects has been the simple frequency glide as studied by Nabelek, Nabelek, and Hirsh (1973).

Lewis, Cowan, and Fairbanks (1940) asked subjects to indicate whether a frequency glide made a larger or smaller frequency transition than the difference between two following tones. They found that the matched width was always smaller than the actual glide excursion and that a short duration glide (100 msec) always matched a smaller interval than a longer duration (300 msec) frequency glide with the same amount of frequency change. In the second half of their study they asked music majors to name the musical interval equivalent to the extremes of frequency glides. The variables studied were the waveform (sinewave or linear based) and the direction (up or down) of the frequency changes in the glides. The direction of the glide had no

effect. Sine waveform glides, however, were consistently perceived as wider than linear glides.

These results agree with the results of a frequency modulation study by H. G. Seashore (1932). Music majors, with ear training, were asked to name the interval corresponding to the extreme pitch excursions of vibrato tones. The actual widths were consistently underestimated, but the estimates increased with increases in the true excursions. Clearly the entire frequency excursion of the vibrato tones was not perceived by the subjects. These experiments show that absolute frequency excursion is only one factor affecting the perceived extent of a frequency change; also important are the rate of change of the frequency and the form, or microstructure, of the frequency change as well.

Some more recent work has focused on this inability to track the actual excursion of frequency modulation and the effect of the modulation waveform. A successful experimental technique has been to directly compare frequency modulation by different modulation waveforms. One of the signals is varied in modulation width until the two modulations produce sensations of equal frequency excursion. If the waveforms had no effect on the tracking of the actual frequency changes then at the point the two FM signals sound equally wide their actual peak frequency excursions should also match. For all waveforms A and B the peak excursions $\pm \Delta f_A$ and $\pm \Delta f_B$, are generally different when the two types of modulation produce equal width sensations. The point of subjective equality (PSE) for the ratio of the two frequency excursions is defined by

$$R_e(A/B) = \Delta f_A / \Delta f_B.$$

Divenyi and Hirsh (1971) compared triangle waveform modulation

(T) and trapezoidal waveform modulation (Z) to a constant width square wave frequency modulation (Q) and found PSEs of

$$R_e(T/Q) = 1.72 \quad \text{and}$$

$$R_e(Z/Q) = 1.07$$

for modulation rates between 4 and 12 Hz. These values indicate that a triangle wave FM signal must have almost twice the frequency excursion of the square wave modulation to sound equally wide. The trapezoid waveform FM excursion need only be slightly larger than the square waveform excursion for the two modulations to sound equally wide.

Hartmann and Long (1976) and Hartmann (1977) compared triangle waveform FM with sine waveform FM (S) under a wide variety of conditions and found that

$$R_e(T/S) = 1.22 \pm .03$$

indicating that the triangle waveform FM must also be wider than sine-wave FM for subjective equality for width. An interesting finding of these studies is that for the range of widths investigated the sine and triangle waveform modulations were indistinguishable. Subjects trained to identify the triangle FM while watching an oscilloscope trace of the modulation waveform could no longer perform better than chance at distinguishing the sine and triangle modulation waveforms when the oscilloscope was removed. More recent informal studies have shown that the width of the frequency excursions must be increased to $\Delta f_s = \pm 40$ Hz before the waveforms could be identified.

Views of Perception. A perceiving individual is constantly receiving vast amounts of information. In order to deal with this great flow the mechanisms of perception must be efficient. Some data must be discarded as irrelevant and the rest must be condensed to a manageable volume.

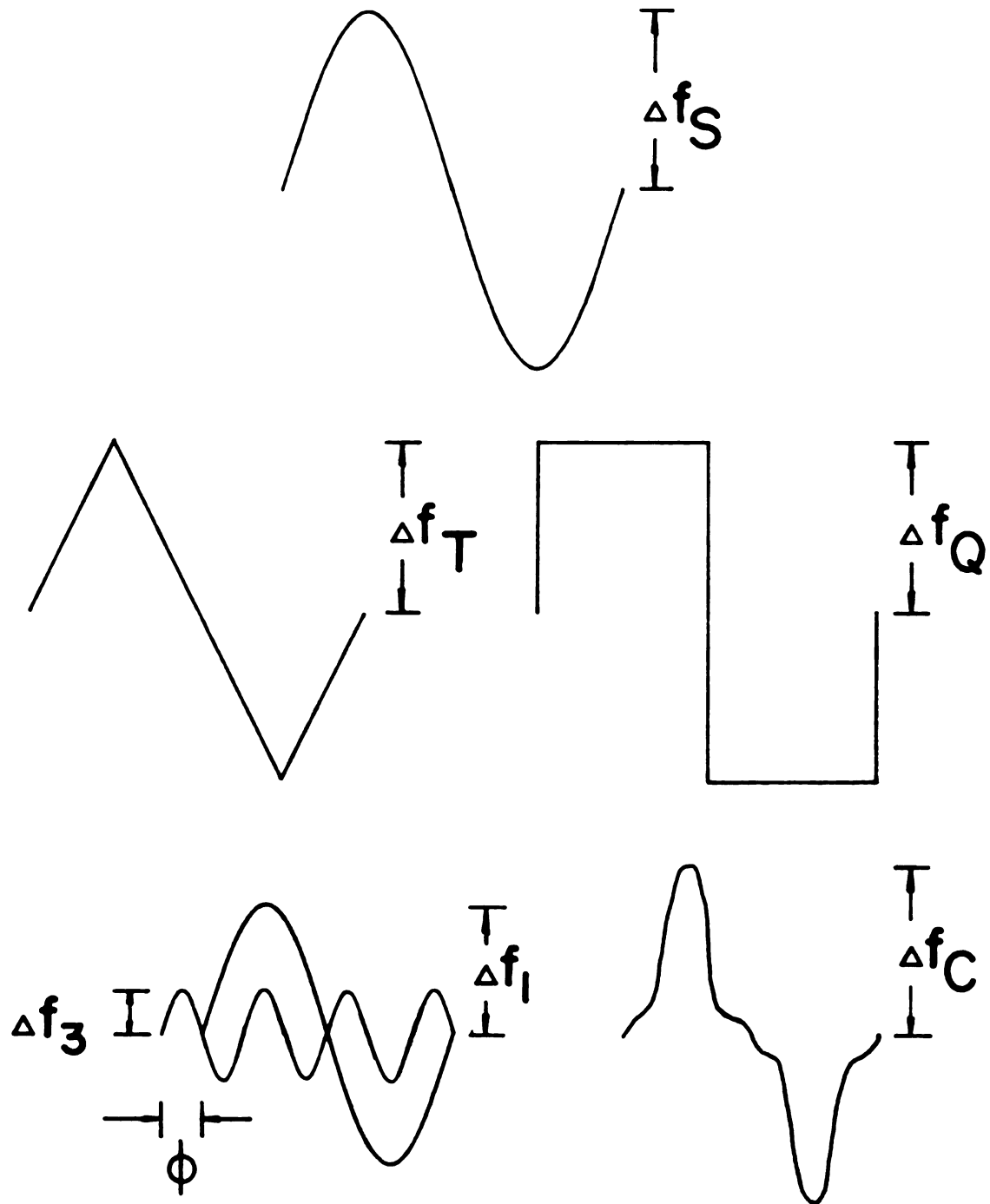


Figure 1. Example waveforms.

The question of how the perceiver decides what to discard and in what way to deal with the remaining information is the central problem of the study of perception. In order for a subject to be able to decide which stimulus had the wider frequency excursion we must assume that information regarding some aspects of the variations of frequency over time was saved. Higher levels of cognitive processing might then assess the saved information to make the final width comparison required by the experiment. Models attempting to explain the perception of frequency modulation width must explain the nature of the process for extracting information from the incoming FM signal. Ultimately, the specific biological sequences involved must be revealed.

At the present we must settle for logical models that are intended only to match the results of the biological processing. These logical models capture the formal properties of the process of extracting information from the stimulus. Good models must accurately predict the responses of the subjects to the stimuli, but they must also succeed at this while using the least amount of information (efficiency). We believe that human perceivers must discard at least part of the information in the stimulus, and while a model utilizing all of the information of the stimulus may be able to accurately predict the subjects' responses it need not closely reflect the type of processing the biological mechanisms perform.

This paper looks at a type of model that represents a "spectral" view of frequency modulation perception. In this approach to FM perception the modulation waveform of the stimulus is analyzed into its Fourier components and their respective amplitudes. The phase relationships between those components are discarded as irrelevant. In general, modulation width is determined by some combination of those

component amplitudes. This singular focus on component amplitudes, totally ignoring phase information, is admittedly a narrow interpretation of the spectral viewpoint.

If, in a spectral model, the phase information was also saved, the entire stimulus would be completely specified by the retained information. This would be equivalent to no reduction of information from the stimulus. Therefore, because of the requirement to abstract information from the input this type of spectral model would be unacceptable.

All other models that might explain the subjects' responses will be grouped together under the title "temporal" models. These models attempt to explain the responses without recourse to the Fourier components of the modulation waveform. Some may require less information be stored about the stimuli than specified spectrally and some may require more. In general, these temporal models will remain unspecified. In these temporal models the form of the incoming signal is processed and represented internally through an isomorphic mapping of the retained temporal structure of the stimulus into its internal representation.

Models. The simplest model for comparing the widths of FM signals is to compare the absolute frequency excursions of the modulation signals. This is a temporal model. The single prediction from this model is

$$R_e(A/B) = 1.0$$

for all waveforms. None of the above studies found that value of R_e to hold.

Hartmann (1978) outlines three possible spectral models that might explain the data. The first is based on the work of Kay and Matthews (1972) who investigated detection of FM for varying modulation rates.

They found that modulation width detection threshold for FM at a specific modulation rate was increased by preexposure to superthreshold FM at the same rate. This led them to the notion of multiple tuned channels for specific FM frequencies about an octave wide, much like the channels described by Blakemore and Campbell (1963) for visual spatial frequency. If this idea of tuned channels is correct its extension to determining modulation width sensation might be as follows: the auditory system performs a Fourier Analysis on the modulation waveform and the amplitude of each frequency component that is above threshold within its own channel becomes available to contribute to the sensation of width. The simplest version of this could be labelled a Fundamental Extractor model. In judging relative width only the Fundamental components of the two modulation waveforms are compared. When the two fundamental components are of equal amplitude there is a sensation of equal width. The combination rule is to discard the information from all components other than the fundamental. The predictions of this model are shown in row 2 of Table 1. The model predictions for sine-triangle modulation comparisons are very close to the data. The triangle versus square waveform modulation prediction clearly fails to come close to the data. The trapezoid versus square modulation waveform predictions, however, are not grossly wrong.

The second spectral model is labelled the Root Mean Square (RMS) model. This is a combination rule for the harmonic components given by:

$$RMS_{FM} = \frac{1}{\sqrt{2}} \left(\sum_{n=1}^N a_n^2 \right)^{\frac{1}{2}} .$$

This form yields a single value for each modulation waveform and any two stimuli with the same RMS_{FM} values should sound equally wide.

Table 1. Predicted and observed R_e -values.

	$R_e(T/S)$	$R_e(T/Q)$	$R_e(Z/Q)$
Predicted R_e			
Model-Peak	1.00	1.00	1.00
FE	1.23	1.57	1.11
RMS	1.23	1.73	1.22
Observed	1.22_1	1.72_2	1.07_2

1. From Hartmann and Long (1976) and Hartmann (1977)
2. From Divenyi and Hirsh (1971)

In the case of the sinewave modulation compared to triangle modulation, column 1, the RMS prediction is in excellent agreement with the data. For square wave modulation compared to triangle modulation the RMS prediction is also very close to the data. For square versus trapezoid modulation the prediction is not so good and clearly fails. The success of the above two models clearly warrants their further investigation.

The final model outlined by Hartmann (1978) is roughly equivalent to a low-pass filter system. The combination rule would be to differentially weight the component amplitudes with the lowest components receiving the most weight. Hartmann (1978) chose, however, to view this model temporally by referring to its equivalent: the Lag-Processor. In this interpretation the system "views" the stimuli through an exponentially decaying time window. The further in the past the actual frequency deviation occurred the less weight it is given for determining modulation width at the present instant. The output of the system is then evaluated for a peak that determines the modulation width. Hartmann assumed that the time constant of the window was fixed. Because of this the model predicted changes in R_e with changes in the modulation rate. Hartmann (1977) found no such changes and so the model was ruled out with no further analysis.

The experiments described below are an attempt to specifically test the spectral models. The presence of phase effects is the critical test.

To facilitate comparisons to the previously mentioned studies the data from Experiment I are first presented using the R_e -values obtained. This leads, however, to a graphical presentation that appears to show a significant phase effect. To correct for this the

data are replotted using a different measure, Q_e . Data from all subsequent experiments are described using the new measure, Q_e .

Experiment I

Work by Graham and Nachmias (1971) points to a paradigm that seems especially suitable to the study of FM width perception. Graham and Nachmias were testing single versus multiple channel models for visual spatial frequency and chose a stimulus that had just two Fourier components. This allowed access to two independent channels simultaneously using only three parameters: the amplitude of each component and the phase relationship between them. An analogous set of stimuli may be produced in the auditory domain. Specifically, an FM signal in which the frequency modulation waveform has only two spectral components: a fundamental component and its third harmonic. This choice of components satisfies two conditions: they are more than an octave apart and so not in the same channel for FM frequencies, if such channels exist (1/3 octave channels; Kay and Matthews 1972), and they are consistent with the previous studies using spectrally complex waveforms in that triangle, trapezoid, and square waveforms only have odd harmonics.

These complex modulation waveform FM signals may be compared to sine waveform modulation FM to measure the Point of Subjective Equality (PSE) for modulation width.

The Fundamental Extractor model predicts that the two modulation waveforms will yield equal width sensations whenever the fundamental component of the complex modulation waveform and the sine modulation waveform have equal amplitudes ($\Delta f_s = \Delta f_1$). This relationship should hold for all phase relationships between the complex modulation

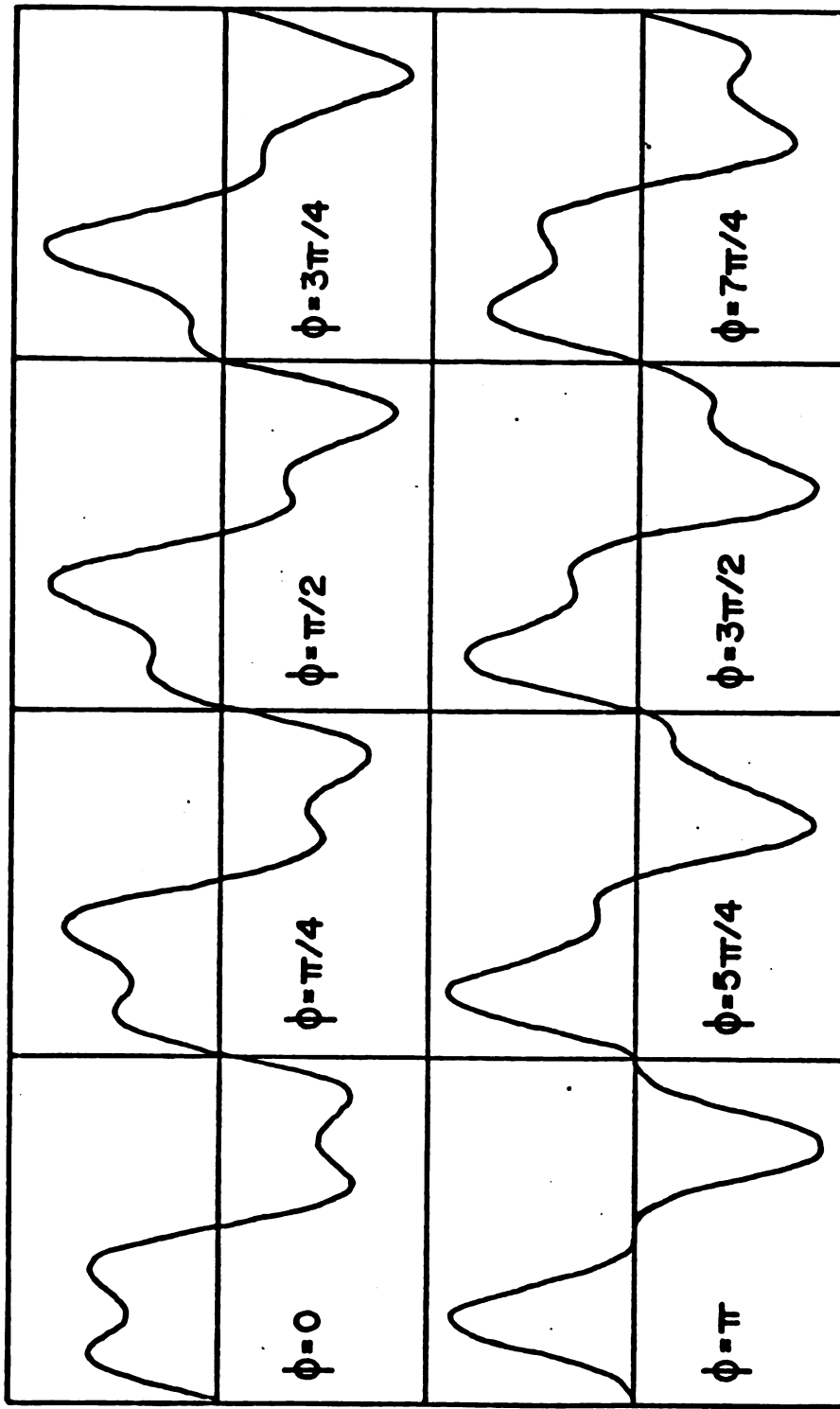


Figure 2. Complex waveforms.

components (ϕ) and for all levels of overall size. The RMS model predicts equal width sensations whenever the RMS value of the complex modulation waveform equals the RMS value of the sine modulation wave. This prediction is also independent of phase and overall width.

The variation of the value of ϕ thus serves to test the spectral models. If changes in the measured PSE are found as the value of ϕ changes the proposed spectral models may be rejected. Temporal models will have to be found to explain the data.

Method.

Subjects. Subjects were volunteers selected for their ability to match pitches correctly and quickly in a Method of Adjustment screening task. These subjects also had to be available for the extended length of time necessary to collect psychophysical-type data. These subjects were paid for their time. The author also scored well on the screening task and participated as a subject. Three subjects provided the bulk of the data. Three other subjects were tested using a subset of the parameters. These other subjects served to verify the responses of the three main subjects.

Procedure. The relative perceived frequency modulation widths produced by the two component complex modulation waveforms were measured by finding the PSE for modulation width between the complex waveform FM and sine waveform FM. The sine modulation waveform has only one Fourier component and always matched the modulation rate of the fundamental component of the complex FM modulation waveform.

To measure the PSEs the amplitude of the sine waveform FM (Δf_s) was kept constant while the amplitude of the complex waveform (Δf_c)

was varied using a simple single up-down staircase procedure. The subject was asked to choose which of the two FM tones had the wider modulation for each trial. If the complex FM was judged to be wider, on the next trial the width of the complex FM (Δf_c) was decreased by about 3%. If the standard, wine waveform FM was judged to be wider, on the next trial the width of the complex FM was increased by about 3%. The starting point of the staircase was chosen arbitrarily and the staircase was terminated after the subject reversed the direction of the staircase 26 times (usually about 50 trials). The values of the first four reversal points were discarded and the arithmetic mean of the last 22 was used as the estimate of PSE. Each subject performed a minimum of three staircase runs at each parameter set and the mean of these multiple PSEs was taken as the subjects' final PSE.

The width of the standard FM, Δf_s , and the waveform of the complex FM as specified by the phase relationship between the two components, ϕ , were the variables studied. PSEs were collected in random order for each of the three values of the excursion of the standard, $\Delta f_s = +2.5$ Hz, $+5$ Hz, and $+10$ Hz, completely crossed with eight values of the phase relationship in the complex FM, $\phi = 0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2$, and $7\pi/4$. This generated 24 PSE data points for each of the three subjects. The three additional subjects were only given two values of the standard excursion, $\Delta f_s = 2.5$ Hz and 10 Hz, crossed with four values of the complex phase relationship, $\phi = 0, \pi/2, \pi$, and $3\pi/2$.

Stimuli. The complex modulation waveforms used are illustrated in Figure 2. The instantaneous frequency produced by these modulation waveforms is given by

$$f_c(t) = f_o + \Delta f_1 [\sin(2\pi f_m t + \theta) + 1/3 \sin(6\pi f_m t + \theta + \phi)].$$

Angle θ is a common phase angle chosen so that $f_c(t = 0) = f_o$ (as for the sine wave).

The amplitude of the third harmonic component of this modulation waveform was fixed at 1/3 the amplitude of the fundamental component. This matches the amplitude of the third harmonic of a square wave. By this choice, the third harmonic component is kept at a reasonable level relative to the fundamental and maintains a visually noticeable effect on the oscilloscope trace of the modulation waveform.

The instantaneous frequency produced by the sine modulation waveform is given by the following

$$f_s(t) = f_o + 4f_s \sin(2\pi f_m t).$$

The modulation rate f_m , was chosen to be 4 Hz. This rate was frequently chosen by the researchers mentioned above. Their choice of 4 Hz was probably determined by data of Zwicker (1952) showing subjects to be most sensitive to FM at that rate.

The carrier frequency was about 800 Hz, but was varied from trial to trial within the range of 775 to 825 Hz. This served to decrease any effects due to local variations in sensitivity along the basilar membrane, (Van den Brink, 1971, 1972).

Stimulus Production. The stimuli for all experiments were frequency modulated sinewave signals. The sinewave carrier signal was produced by a Wavetek Model 116 Voltage Controlled Oscillator (VCO). The frequency of the carrier was varied by control voltages from a microcomputer, (Klein, Gable Edmunds, Eicher, & Hartmann, 1978). The FM waveforms were calculated by the microcomputer and stored in memory. The memory structure for those stored waveforms was a 256 sample, 16-bit signed integer vector holding values representing the frequency deviations for a single cycle of the FM waveform. The

modulation rate, f_m , was determined by the rate at which the computer stepped through the vector, resetting the VCO control voltage according to the current element of the vector. For example, a 4 Hz FM could be produced by repeatedly cycling through all 256 elements of the vector once each 250 msec, presenting each element for 976 microseconds. An 8 Hz FM would require a 488 microsecond sample time. The duration of the FM signal was determined by the number of times the "playing" of the vector was repeated. The amplitude of the signal was controlled by a rectangular envelope produced by an Exponential Voltage Controlled Amplifier (EVCA) also controlled by the microcomputer. The calibration frequency (800 Hz) for the carrier was verified using a frequency counter and the calibration amplitude (75, dB SPL) was measured electronically using a Ballantine Model 320A voltmeter measuring True-RMS. Subjects listened to tones with headphones (Beyer DT-48) while sitting in a sound proof room.

Trial Structure. A trial consisted of two 2-second long tones presented in a Two-Interval Forced-Choice (2IFC) paradigm at 75 dB SPL and separated by 250 msec. Both tones had the same carrier frequency and modulation rate. One tone was always the standard sine waveform FM tone with a fixed frequency excursion, Δf_s . The standard FM signal was presented first with probability $p = .5$. The standard excursion, Δf_s , remained constant within a single experimental run, but was varied as a parameter.

A green warning light lasting 250 msec would indicate the start of a trial. The two stimulus intervals would immediately follow and a different color light was turned on simultaneously with each interval to indicate which button to push to choose that interval. When the stimuli were finished a red light came on to request a

response. The next trial began 750 msec after a response was made, see Figure 3. When the criteria for termination of an experimental run was met all lights were turned on simultaneously to notify the subject.

Results-Discussion. The results of Experiment I are shown in figures 4 through 8. The data shown in each graph represents values of R_e collected at each parameter set. Figure 4 shows the group results for the 3 main subjects. The predictions from each of the 3 models are also included; the peak detection model, the RMS model and the Fundamental Extractor model are shown. The apparent phase dependence of the RMS and Fundamental Extractor model predictions is discussed below.

From the data shown there are two main dependencies, an effect due to the size of the standard, Δf_s , and an effect due to the phase angle between the components of the complex modulation waveform, ϕ . In all instances the peak detection model has obviously failed again. The other two models do not fail as completely, but still appear inadequate.

In the discussion above regarding the two models, the RMS and Fundamental Extractor, it was pointed out that there should be no effect due to phase. The calculated prediction, however, show changes with ϕ that must be explained. The nature of the R-value is the cause of this apparent discrepancy. The R_e -value is calculated from the peak excursion of the waveforms. When adding together the two components of the complex waveform, if the sizes of those components are kept constant and only the phase between them is changed, the peak of the resulting waveform will change. This effect can be seen in Figure 2 showing all of the complex waveforms used. The amplitudes of the two components remained constant while the phase was changed to calculate

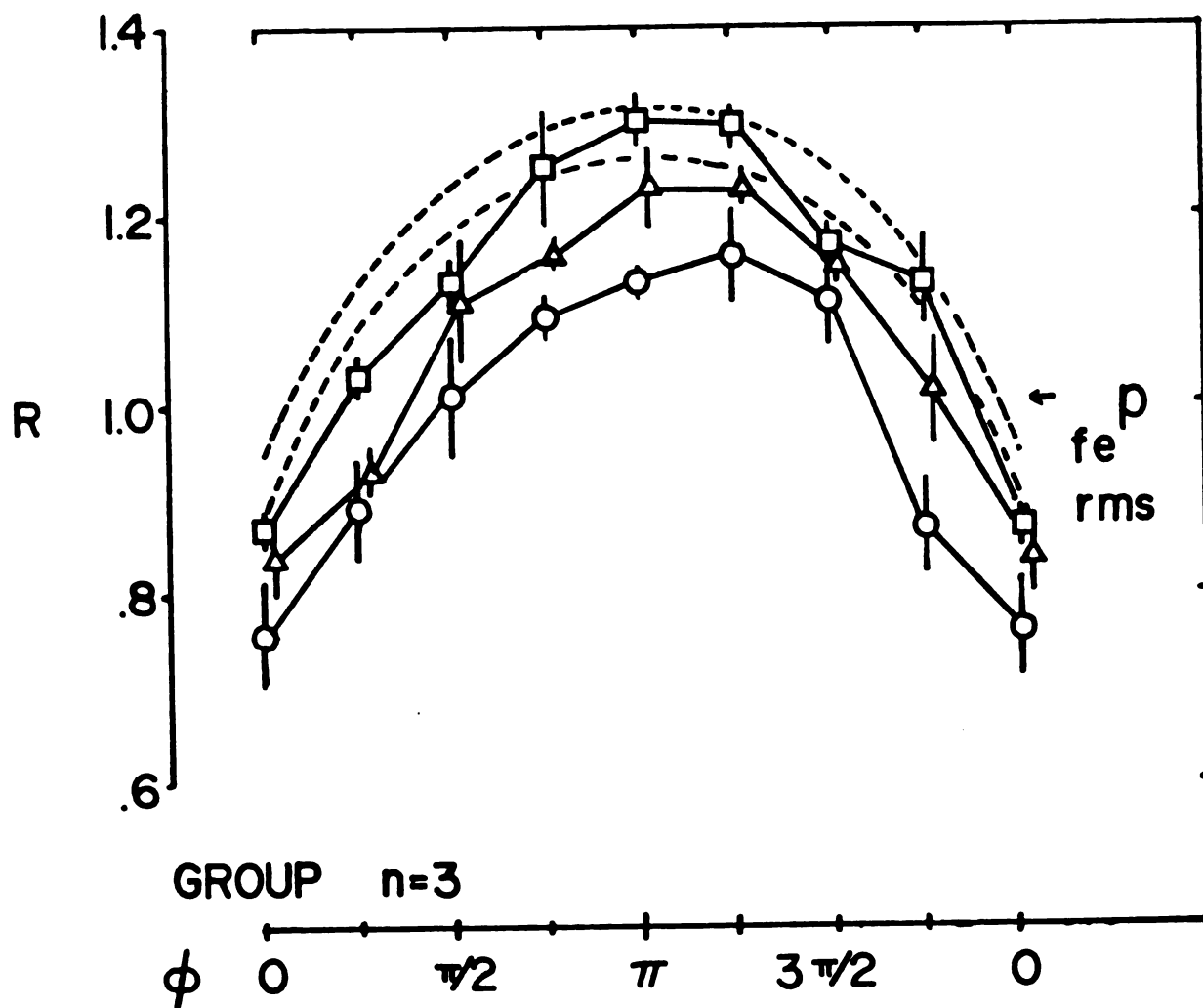


Figure 4. R_e -values. Group data Experiment I.

For this and all following figures \square represents $\Delta f_s = 2.5$ Hz

Δ represents $\Delta f_s = 5$ Hz and \circ represents $\Delta f_s = 10$ Hz.

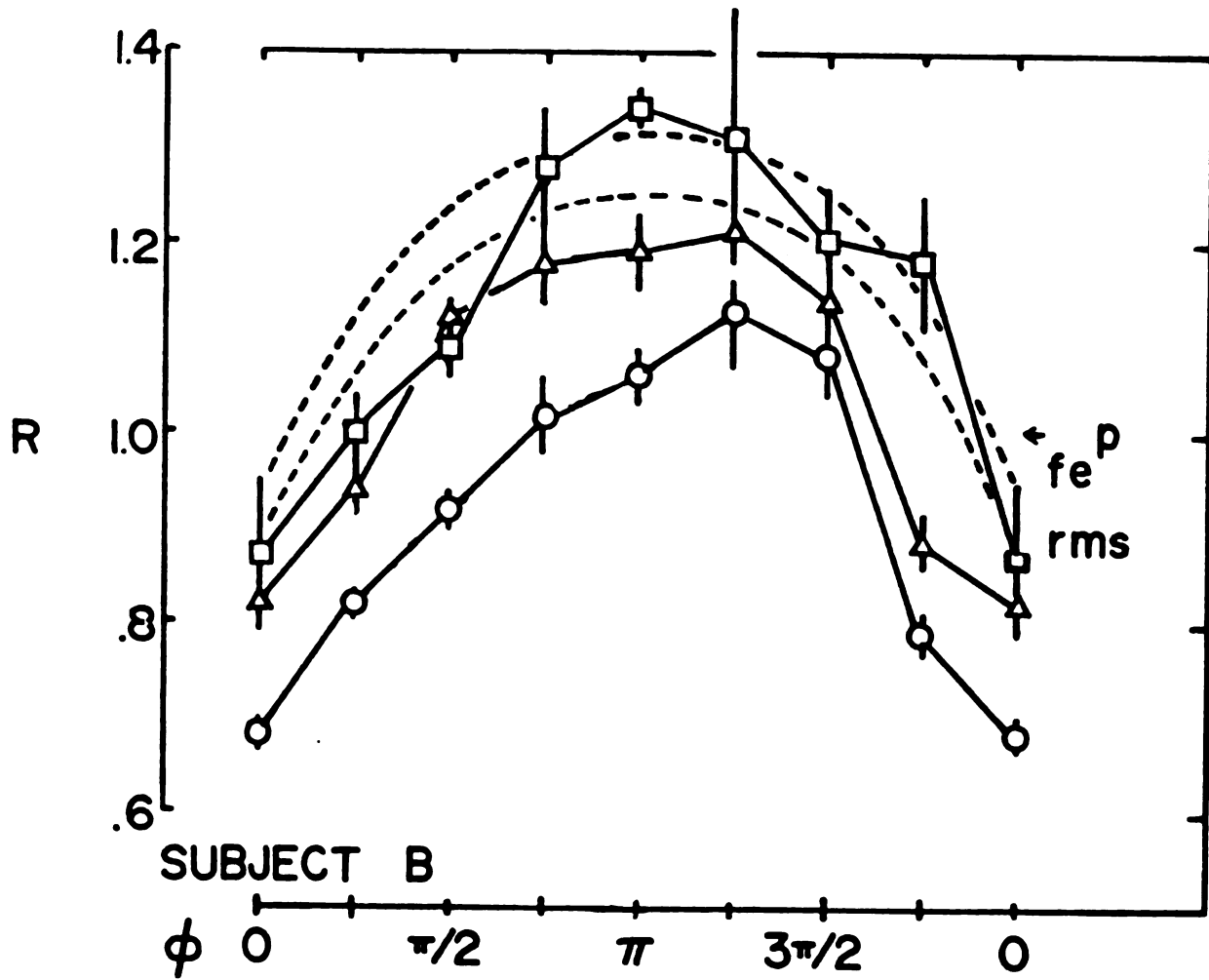


Figure 5. R_e -values. Subject B Experiment I.

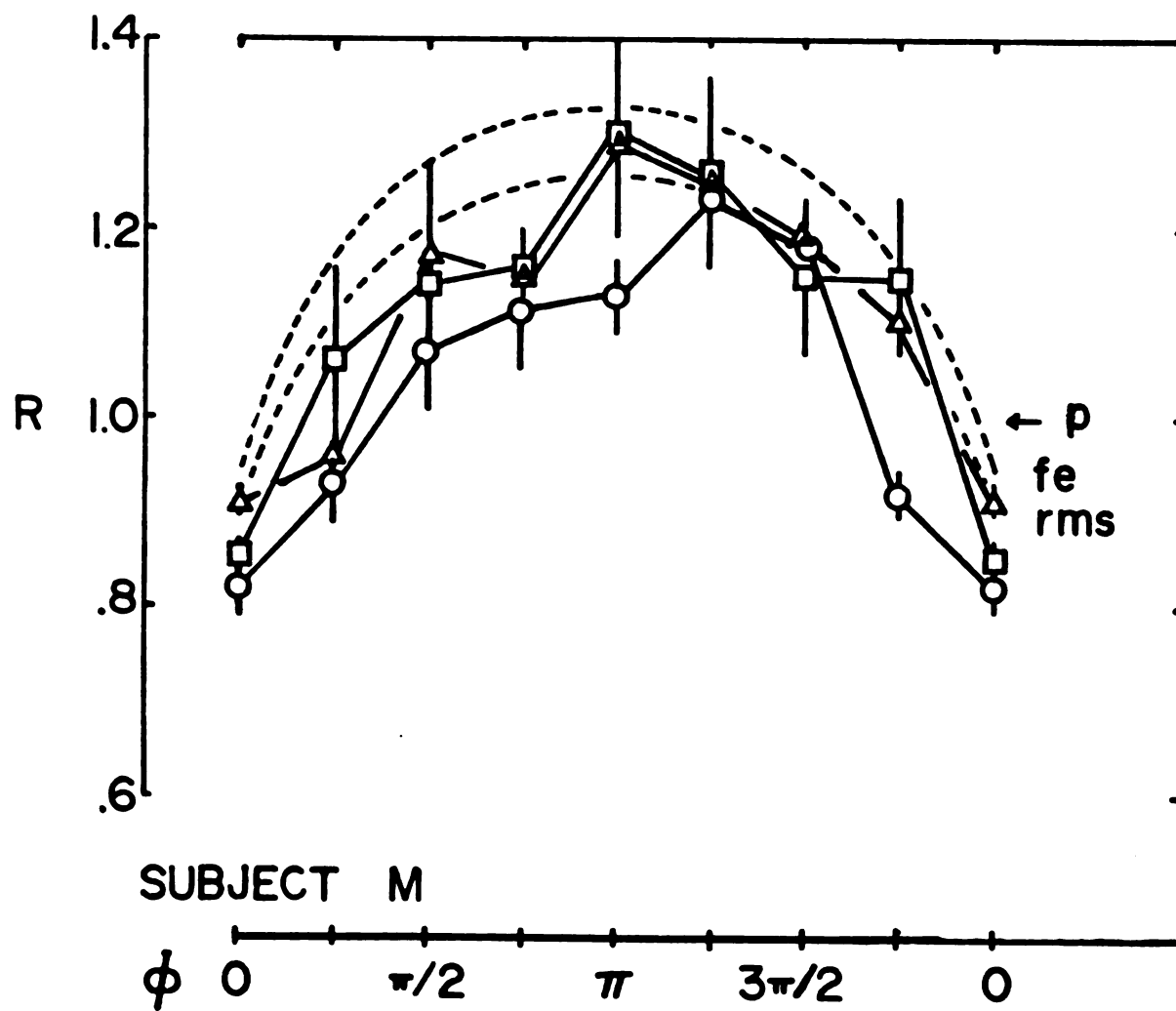


Figure 6. R_e -values. Subject M Experiment I.

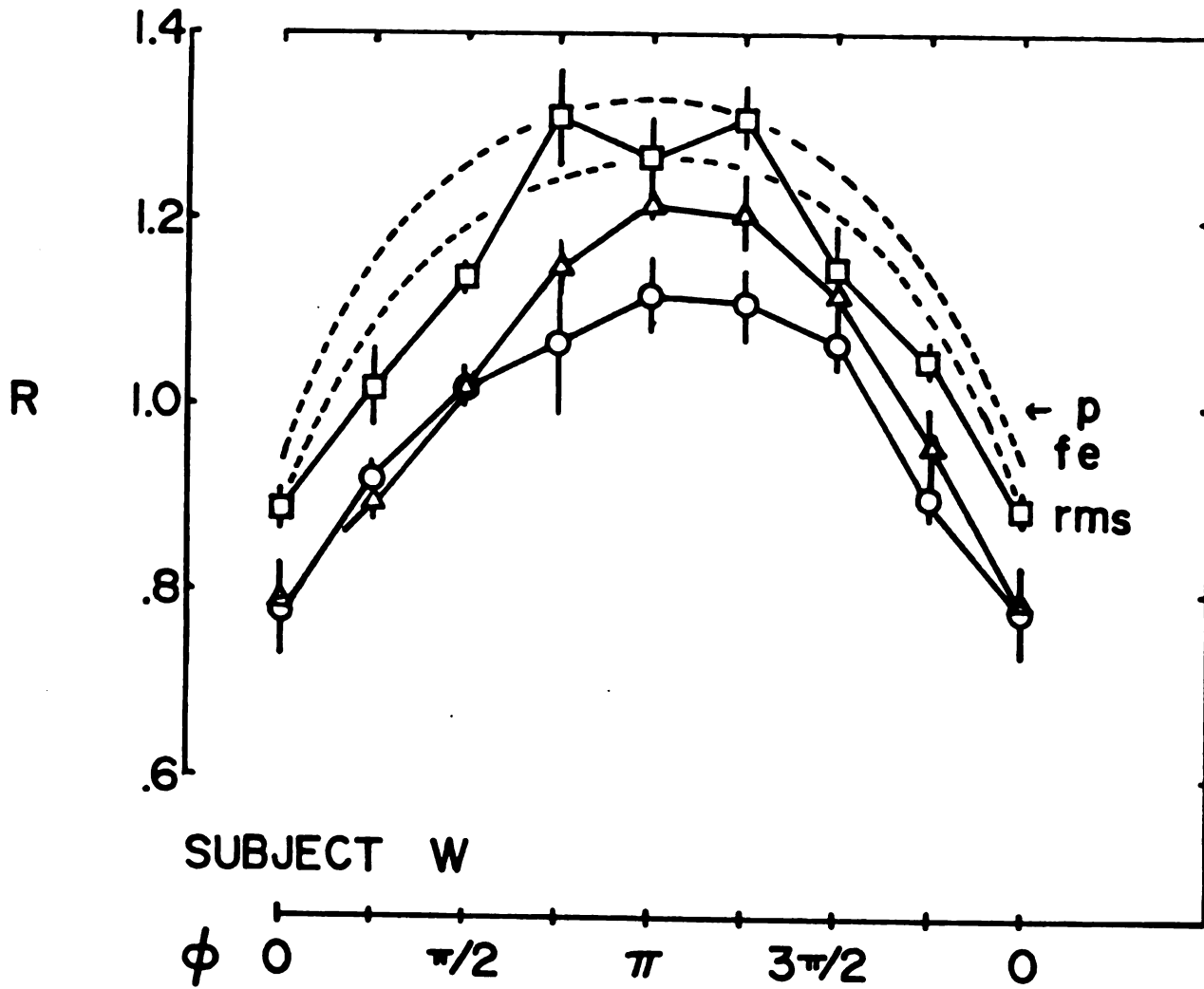


Figure 7. R_e -values. Subject W Experiment I.

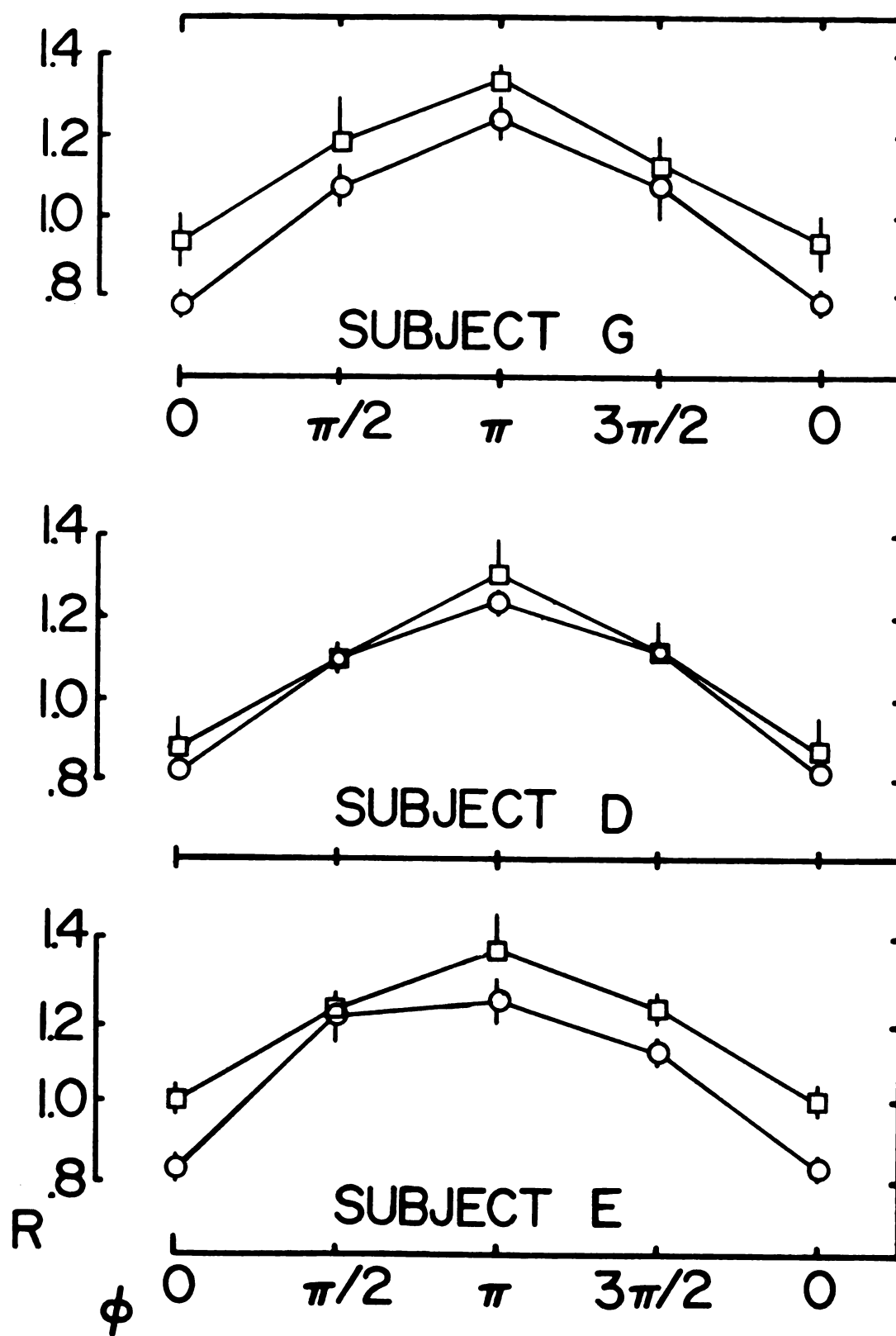


Figure 8. R_e -values. Auxillary subjects Experiment I.

and plot the various waveforms. For the predictions of the Fundamental Extractor model, so long as the fundamental component has an amplitude equal to the sinewave standard's amplitude the signals should produce sensations of equal width. The predicted R_e -values will vary with the changes of the peak even while the fundamental remains constant. The predicted values of the RMS model show the effect based on the same type of analysis. Because the amplitudes of two components remain constant the RMS value for the complex modulation also remains constant.

The graphs of the data show that the measured R_e -values roughly follow the curves of both of the models. It is difficult, however, to determine the extent of the deviations from the predicted values because of the curve. To better illustrate the actual phase independence and observe the deviations from the predicted values the predictions were recalculated and the data replotted in figures 9 through 13 using the ratio of the amplitude of the fundamental component of the complex modulation waveform, Δf_1 , to the amplitude of sine waveform, Δf_s . Call this Q_e where

$$Q_e (C/S) = \frac{\Delta f_1}{\Delta f_s}$$

The predicted Q_e -values for the two models are now constants. The gross curve in the data has been eliminated. It is easier to observe deviations from the predicted values. Two main effects are still evident, the effect due to Δf_s and now an effect in which the Q_e -values for ϕ between 0 and π were lower than the Q_e -values for ϕ between π and 2π (i.e. $\phi = 0$), an asymmetry centered at $\phi = \pi$. This effect showed up quite strongly in the group data. Data for individual subjects showed the same effect to varied extents. To test this observation statistically 95% confidence intervals were calculated

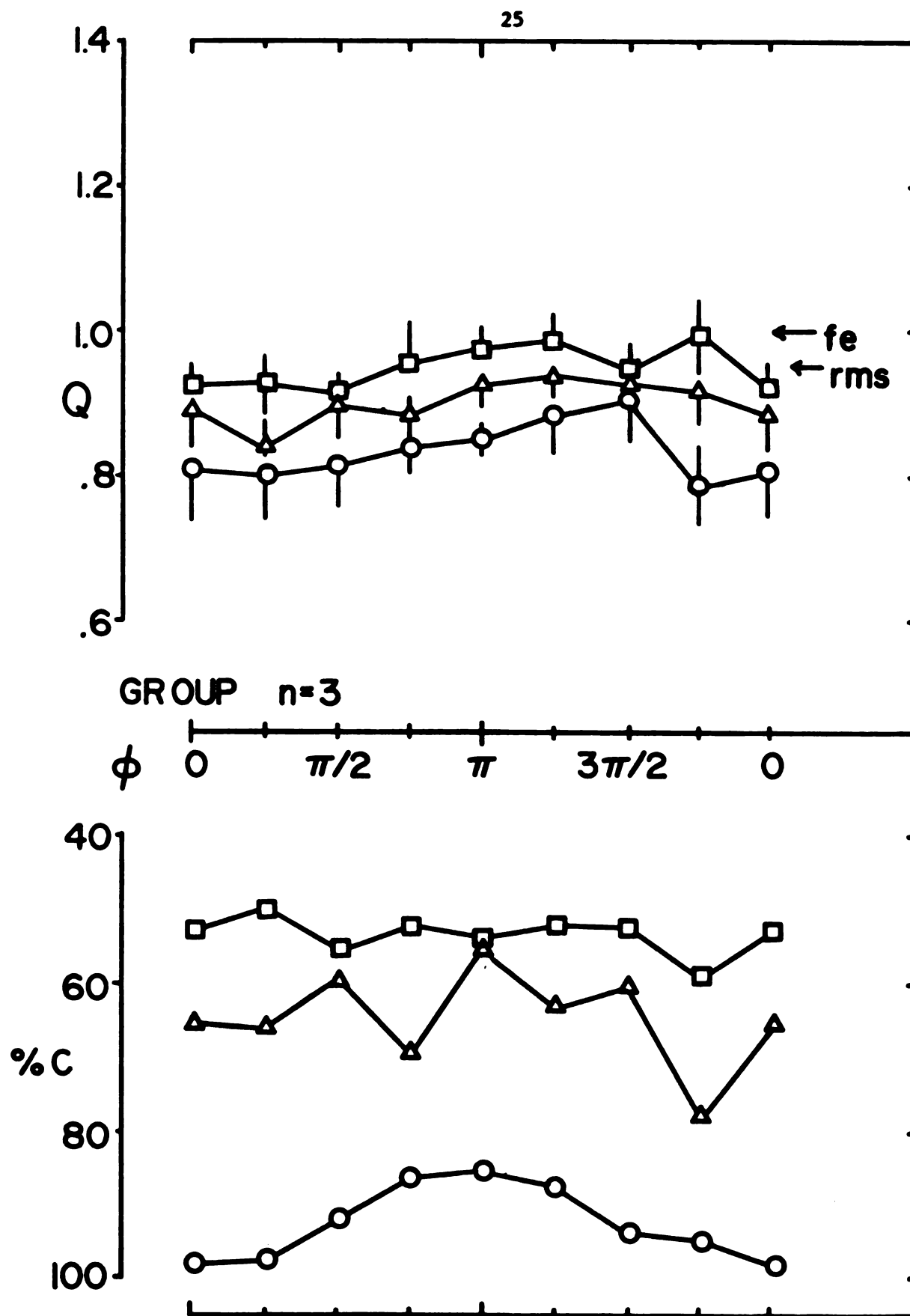


Figure 9. Q_e -values. Group data Experiment I and II.

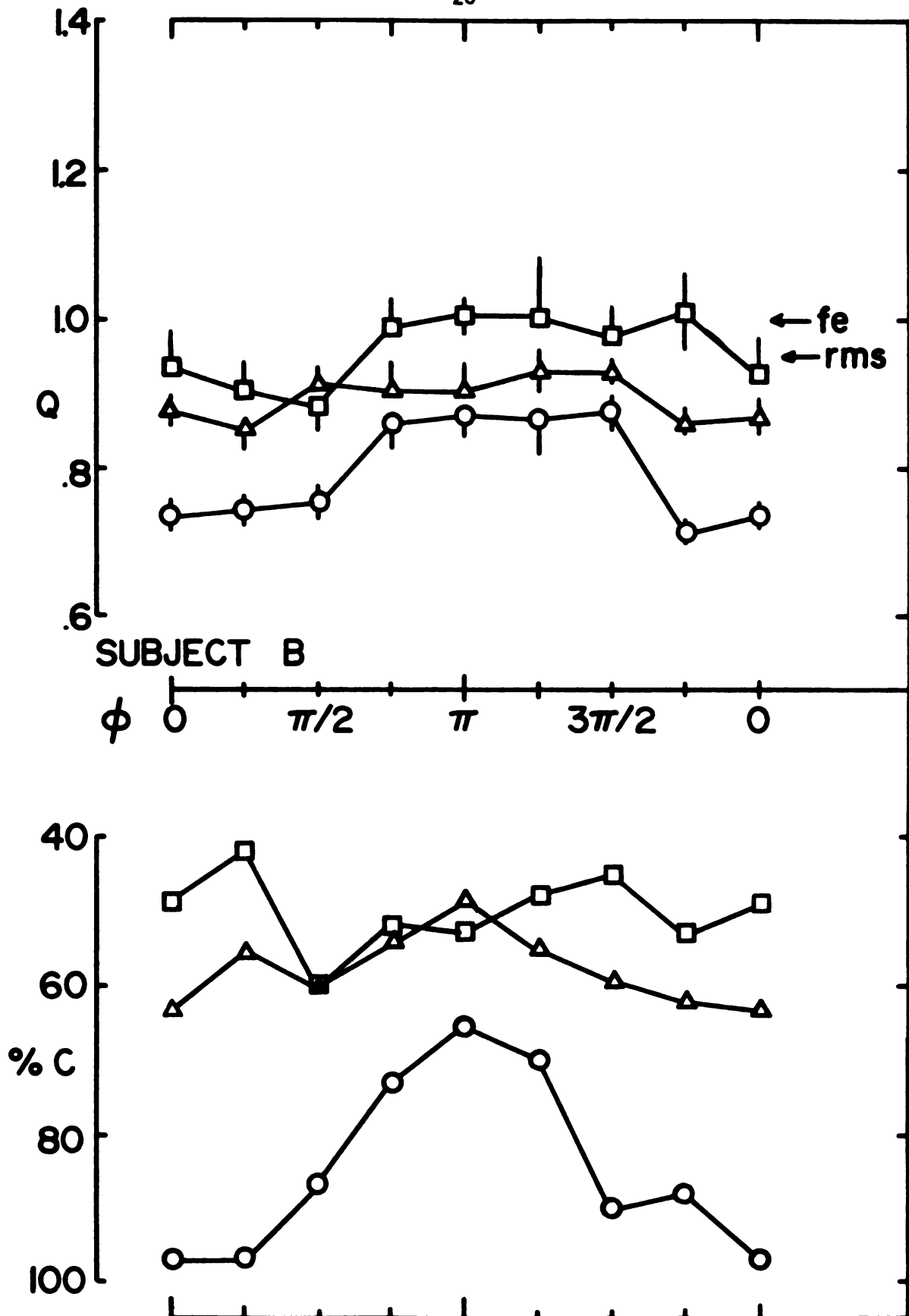


Figure 10. Q_e -values. Subject B Experiment I and II.

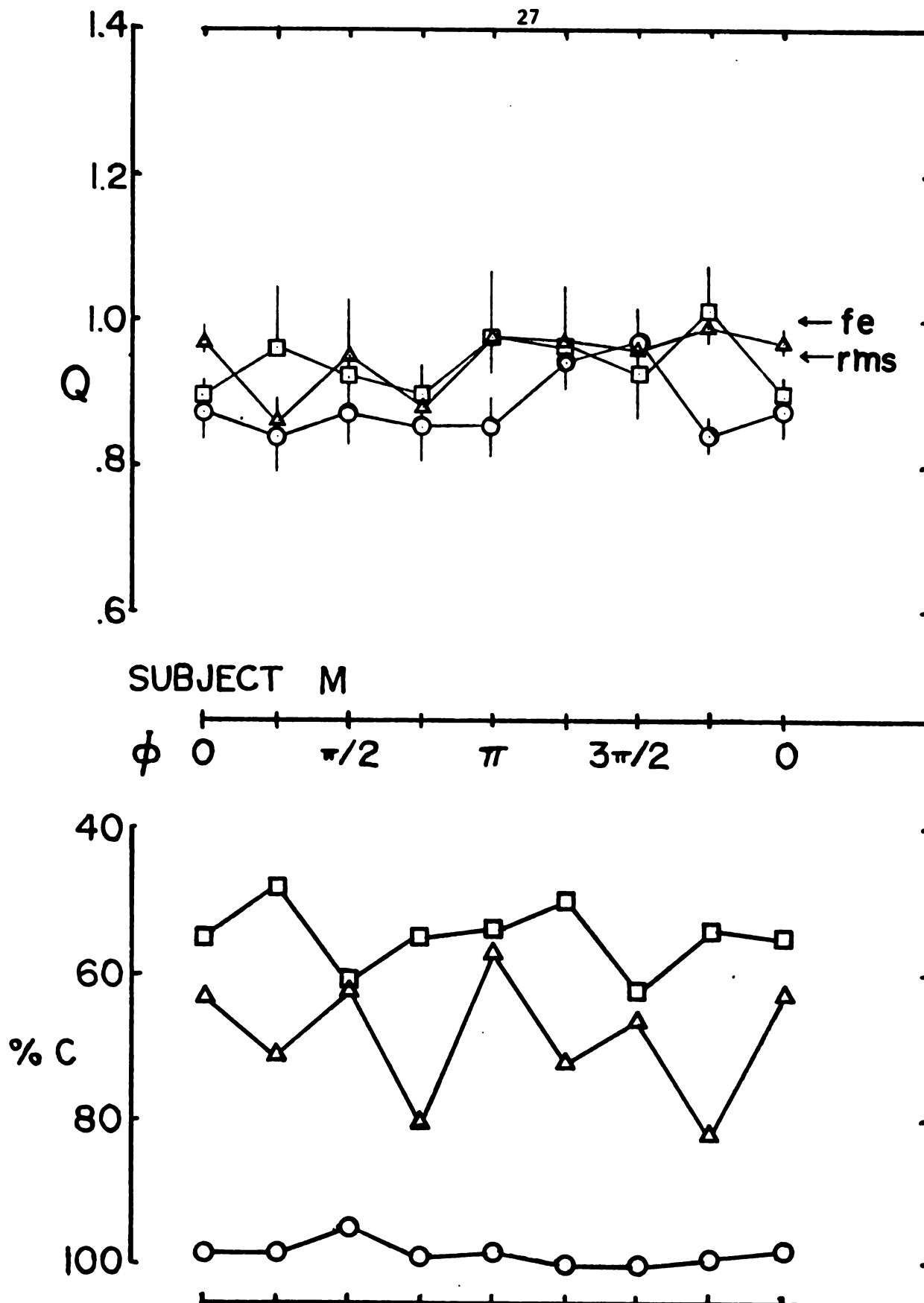


Figure 11. Q_e -values. Subject M Experiment I and II.

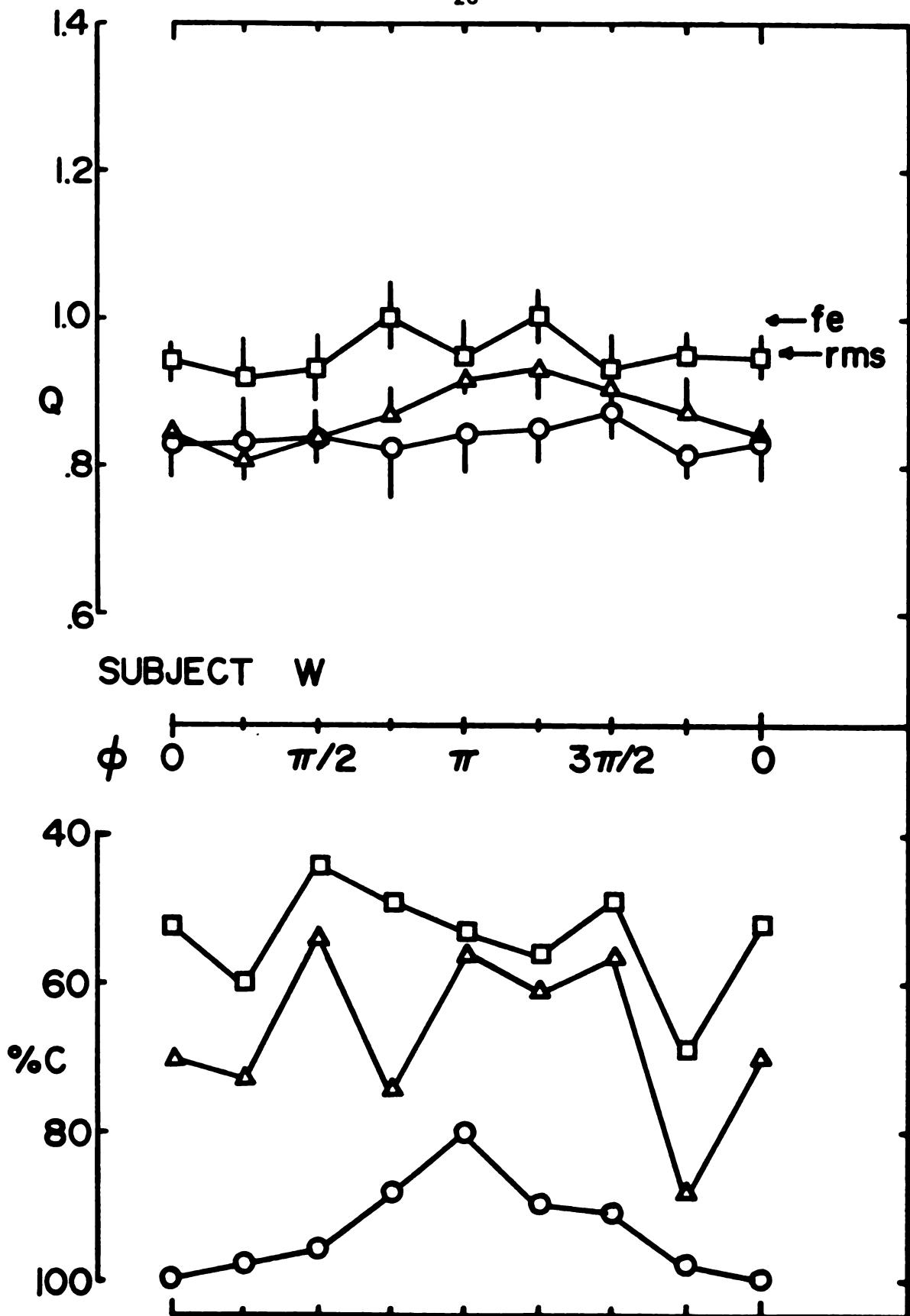


Figure 12. Q_e -values. Subject W Experiment I and II.

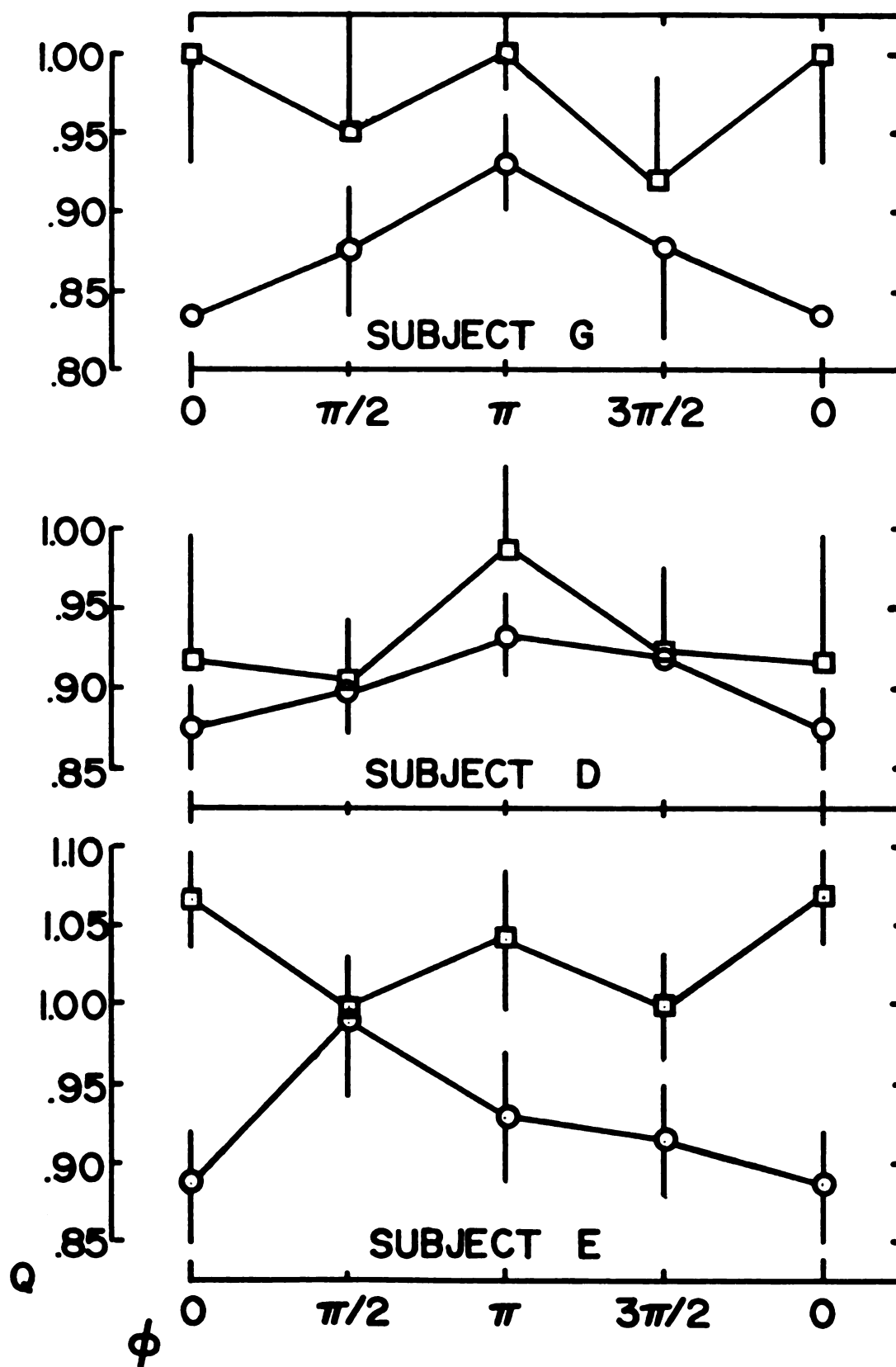


Figure 13. Q_e -values. Auxillary subjects Experiment I.

for the difference between the mean of pooled data from $\phi = 5\pi/4, 3\pi/2$, and $7\pi/4$ compared to pooled data from $\phi = \pi/4, \pi/2$, and $3\pi/4$ for each subject. Table 2. shows these confidence intervals. These confidence intervals are based on the 95th percentile points of the t-distribution. Thus if a confidence interval does not include 0.0 then a hypothesis that the observed difference equals 0.0 must be rejected. From these confidence intervals it can be seen that Subject B only showed a statistically significant asymmetry for data at $\Delta f_s = \pm 2.5$ Hz. Both Subjects M and W only showed significant asymmetry at $\Delta f_s = \pm 5$ Hz. Other asymmetries seen in the graphs are probably not significant due to variance in the data. Subjects B and M do show the asymmetry when data from all values of Δf_s are pooled. The data pooled for all subjects shows significant asymmetry at $\Delta f_s = \pm 2.5$ and ± 5 Hz. Again, it is the variance of the data at $\Delta f_s = \pm 10$ Hz that probably prevents statistical significance. This asymmetry is an effect that may be predicted by the temporal interpretation of the Lag-Processor model discarded by Hartmann (1978). Neither of the other two models predict this. It is this effect that provides the motivation for Experiment IV.

The other major effect that is evident in both the R_e and Q_e transformations of the data is the change toward higher values as Δf_s gets smaller. This is clearly visible in the data for Subjects B and W. Table 3. shows 95% confidence intervals for the differences between the data at each value of Δf_s grouped across values of ϕ .

Positive differences here indicate an increase in the average Q_e -value as Δf_s decreases. Only Subject M showed a non-significant increase and only between data at $\Delta f_s = \pm 2.5$ and ± 5 Hz. It should

Table 2. 95% confidence intervals for differences between pooled data for $\phi = \pi/4, \pi/2, \text{ \& } 3\pi/4$ versus pooled data for $\phi = 5\pi/4, 3\pi/2 \text{ \& } 7\pi/4$.

Subject	$\Delta f_s =$				Overall
	2.5 Hz	5 Hz	10 Hz		
B	.003 < δ < .169	-.010 < δ < .060	-.340 < δ < .114	.007 < δ < .113	
M	-.019 < δ < .151	.937 < δ < .113	-.001 < δ < .113	.927 < δ < .102	
W	-.029 < δ < .057	.023 < δ < .095	-.027 < δ < .055	-.005 < δ < .063	
Group	.016 < δ < .094	.029 < δ < .079	-.001 < δ < .073	.024 < δ < .072	

$$\begin{aligned} \delta &= Q_{e2} - Q_{e1} \\ Q_{e1} &= 1/3 (Q_e(\phi=\pi/4) + Q_e(\phi=\pi/2) + Q_e(\phi=3\pi/4)) \\ Q_{e2} &= 1/3 (Q_e(\phi=5\pi/4) + Q_e(\phi=3\pi/2) + Q_e(\phi=7\pi/4)) \end{aligned}$$

Table 3. 95% confidence intervals for differences between
 Q_e -values across levels of Δf_s .

		$\delta = Q_e(\Delta f_s = +2.5) - Q_e(\Delta f_s = +5)$	$\delta = Q_e(\Delta f_s = +5) - Q_e(\Delta f_s = +10)$
Subject			
B		.020 < δ < .122	.038 < δ < .141
M		-.935 < δ < .046	.018 < δ < .120
W		.052 < δ < .107	.019 < δ < .064
Group		.028 < δ < .076	.040 < δ < .088
		$Q_e(\Delta f_s) = \frac{1}{8} \sum_{n=0}^7 Q_e(\phi = n/4; \Delta f_s)$	

also be noted that as the Q_e -values increase they approach the predicted values of the two main models. Neither of the two models predict this Δf_s effect. Clearly simple versions of the RMS and Fundamental Extractor models are inadequate. The larger the overall width of the stimuli the less accurate these models become. This effect motivates Experiments II and III.

At this point it seems wise to stop and summarize the preceding results before examining their implications for the question of spectral versus temporal processing.

- 1.) The fine structure of the modulation waveform as determined by ϕ has a significant effect on the Q_e -values. For ϕ between π and 2π the Q_e -values are increased.
- 1a.) The asymmetry appears to be largest at larger values of Δf_s , but is not significant at $\Delta f_s = \pm 10$ Hz due to increased variance of the data collected.
- 1b.) Purely spectral models cannot adequately explain this effect. This effect is not large, however, and some spectral model may serve to approximate the data.
- 2a.) The overall size of the waveforms as determined by Δf_s has a significant effect on the Q_e -values. As Δf_s decreases the Q_e -values increase.
- 2b.) In the limit of small Δf_s the RMS and Fundamental Extractor models fit the data approximately.

Only a model taking phase into account can predict the detailed phase-dependent structure of the data. Combination rules for spectral components alone cannot handle these effects. This provides evidence for a temporally-based view of this process. The fine structure of the

modulation waveform is processed and affects the perceived width of the modulation. For a temporal model the rate, the magnitude, and the form of the change all become possible parameters for determining the perceived width. All of these parameters are affected by the changing ϕ . As ϕ changes the peak excursion changes as described above, the rate of change of frequency changes for different segments of the waveform, and the form changes. All these changes are seen in Figure 2. This confounding of independent parameters makes it impossible to specify exactly what is the basis for the phase effect.

The Δf_g dependence does not apply as clearly to the temporal versus spectral processing question. Within the spectral model it is clear that each component affecting the perceived width must be above threshold for detection. As Δf_g increases the size of the complex modulation waveform increases to find a new PSE. This then has the effect of increasing the absolute amplitude of both of the components of the complex modulation waveform. As these components increase in size it is quite reasonable that perhaps the third harmonic makes the transition from below detection level to being well above threshold. If this is the case, then at $\Delta f_g = +2.5$ Hz, only the fundamental component is available to contribute to the width judgement. The Fundamental Extractor model predictions would match the data regardless of whether the real process involved more than one harmonic or not. The fact that the Q_e -values do change as Δf_g increases is evidence against the Fundamental Extractor model, but does not eliminate some other combination rule for the contribution of the modulation waveforms to the perceived width.

For the other spectral model, RMS, it is clear that as the amplitude

of the third harmonic increases to exceed threshold the "perceived" RMS value should more closely approach the actual RMS value of the modulation waveform. The predicted RMS value is based on the actual components of the modulation waveforms. The data show that the perceived relative width deviates from the RMS model predictions as Δf_g increases. This indicates only that RMS is not the appropriate combination rule for the components on the modulation waveforms. It is possible that some other spectral type model may accurately predict these results.

The Δf_g dependence may also be interpreted within the temporal view of perception. It is reasonable to assume that the possible temporal parameters alluded to above have some type of threshold below which they cannot be processed. If this is the case, as Δf_g decreases some of these waveform characteristics may lose some of their ability to affect the perceived modulation width. At large Δf_g many parameters determine the perceived width and at small Δf_g the perceiver must rely upon only those waveform characteristics still above threshold. If none of the characteristics are detectable, no modulation would be perceived at all.

Experiment II

In Experiment I the Δf_g dependence pointed to some process that took into account either the detectability of the third harmonic in the spectral models, (without specifying the component combination rule), or the detectability of the microstructure of the waveforms as necessary for the temporal models. A natural question to arise is 'to what degree are listeners aware of the factors influencing the difference in perceived width between the complex and the sine modulation waveforms?' Can the subjects distinguish between the two modulation waveforms to

the same degree that the perceived relative width differs from some simple model for width perception? Does the discriminability between the waveforms go down as Δf_g decreases? Is there an asymmetry between the discriminability of the waveforms for ϕ on either side of π ?

Experiment II measures the discriminability of the waveforms used in Experiment I. To accurately measure the degree to which the same factor influencing width comparison affects discrimination it is necessary to eliminate as many extraneous cues as possible. It is necessary to equalize the stimuli on as many characteristics as possible. The stimuli used were the same stimuli as were used for Experiment I. That means the stimuli are at the same amplitude and the same center frequency. In order to eliminate perceived width as an identification cue the width specified by the PSEs collected in Experiment I were used. That meant that the only likely cue for discrimination must be the third harmonic and/or its effect on the structure of the modulation waveform.

A spectral view of the discrimination process involves simple detection of the third harmonic component. If it is detected the waveforms will be discriminated. There should be no phase effects in the discrimination data. Any phase effect that might be found would be evidence for rejecting the spectral view.

Method. The task for this experiment was to identify the complex waveform modulation signal when presented two FM signals. One was always sine waveform modulation and the other was complex waveform modulation. The two signals were presented in a 2IFC paradigm where the order of the two signals was randomized. The trial structure was identical to that described above for Experiment I. The modulation excursion, Δf_g

and Δf_c , were chosen according to the group PSE for the three subjects receiving the full stimulus set in Experiment I. The same three subjects participated in this experiment. The group PSE was used so that more than one subject could participate at a time.

Subjects were trained in this task in order to eliminate learning effects. Data reported is from 100 trials collected after it was determined that the subjects had reached asymptotic behavior. The relative widths of the two modulation signals remained constant throughout the trials according to the group PSE. Only the center frequency changed from trial to trial as described above for Experiment I. The percent of correct discriminations for each subject at each value of ϕ and Δf_s was collected.

Results-Discussion. The results of Experiment II are shown in the bottom half of Figures 9 through 12. Figure 9b shows the group data. There was a large effect due to the overall width, Δf_s . As Δf_s decreased there was a decrease in the ability of the subjects to discriminate the two modulation waveforms. The group data at $\Delta f_s = \pm 10$ Hz also showed an effect due to phase. Both subjects B and W showed this effect in their own data while subject M did not. Subject M was clearly too good at this task. Presumably, if Δf_s had been decreased slightly subject M would have shown this phase effect. Unlike the data of Experiment I there is no asymmetry in the group Phase effect.

In general, a comparison between the curves for Experiment I and II shows loose similarities. When pairs of points at a given value of the phase relationship are relatively close together or even reversed in order in Experiment I the same points from Experiment II are also relatively close together. Even some of the idiosyncratic changes in

direction along the curves in Experiment I are paralleled in Experiment II. If it is the case that the detectability of the third harmonic is constant across values of ϕ within a single Δf_g level then those parallels are evidence against the spectral view and in favor of a temporal model. It does show that the same features influencing perceived width also influence waveform discrimination.

The phase effect for the two subjects at $\Delta f_g = \pm 10$ Hz provides evidence that temporal aspects of the modulation waveform are being processed as the subjects perform the discrimination. However, because the stimuli of Experiment II were equalized for width it is not the case that the amplitude of the third harmonic remained constant across values of ϕ . This might mean that the apparent phase effect is due to changes in the levels of the 12 Hz component rather than the changes in the microstructure. Experiment III looks at this possible problem more closely.

Experiment III

In the specification of the modulation waveforms used in Experiment I it is evident that a single element distinguishes the standard sine modulation waveform from the complex modulation waveform; the presence of the third harmonic. Within the spectral view it is the presence or absence of this additional component that initiates the process to modify the perceived FM width. In the temporal view it is not the detection of the third harmonic, but the detection of the changes induced in the modulation waveform by the third harmonic that are important. This distinction is quite subtle, but it is this fact that allows the phase relationship between the two components of the complex modulation waveform to have an effect in the temporal models, but not in the

spectral models.

Obviously, the detection level of the third harmonic should be correlated with the detection level of the induced alterations in the modulation waveform. It is not necessarily the case, however, that the induced alterations are detectable if and only if the third harmonic is above threshold. Specifically, the microstructure of the waveform may affect perception even while the third harmonic is below threshold or the third harmonic may be well above threshold, but its effect on the microstructure such that it is difficult to distinguish the presence of the alterations in the waveform. Applied to Experiment II, this notion of independent thresholds may be easily tested. With specific values for the detection threshold for the third harmonic component of the modulation waveform predictions can be made regarding the discrimination performance of each subject. Experiment III was the measurement of the detection threshold for 12 Hz sine waveform frequency modulation. The amplitude of the modulation waveform measured in Hertz at threshold was the desired result. The threshold for 4 Hz FM was also measured for completeness.

If the detection level of the third harmonic component of the modulation waveform cannot accurately predict discrimination performance spectral type models can be rejected as inadequate for modelling this task.

Method. The paradigm and trial structure was identical to that described for Experiment I except for the following modifications. The complex waveform modulation signal was replaced by a constant frequency tone having the same frequency as the center frequency of sine waveform modulation. The width of the sine waveform modulation varied in a transformed up-down procedure (Levitt, 1970) that converged on the $p = .707$ point

of the psychometric function for detection of the FM. If the subject identified the interval with the FM twice in a row the width of the FM was decreased by about 1.5% on the next trial. If the subject failed to identify the FM on any trial, the FM was increased in width by about 1.5% on the next trial. The mean of the last 22 reversal points again served as the estimate of the point of convergence for that experimental run. For the purposes of this short experiment the FM width at which the FM was detectable at the $p = .707$ level was considered to be the detection threshold.

Subjects did a minimum of three experimental runs at both 4 Hz and 12 Hz. The mean of these multiple estimates served as the final estimate for each subject. The standard error of those means provides a measure of their variability.

Results. The FM detection levels for each of the three subjects are shown in Table 4. The 12 Hz data are also shown plotted as straight lines in Figure 14 relative to the actual size of the third harmonic for each parameter pair used in Experiment II. As shown, subjects M and W had relatively similar thresholds. Both were lower than the threshold for subject B and between the third harmonic levels for $\Delta f_s = +2.5$ Hz and $\Delta f_s = +5$ Hz. Subject W had the lowest threshold and Subject B the highest, almost equalling the level of the third harmonic for $\Delta f_s = +10$ Hz.

Discussion. With the combined data of Experiments II and III it is possible to look at the question of whether the waveform microstructure has effects independent of the threshold of the third harmonic. This is a direct test of spectral type models. If the level of the third harmonic component relative to the individuals' threshold predicts the discrimination performance it will not be possible to reject the spectral type

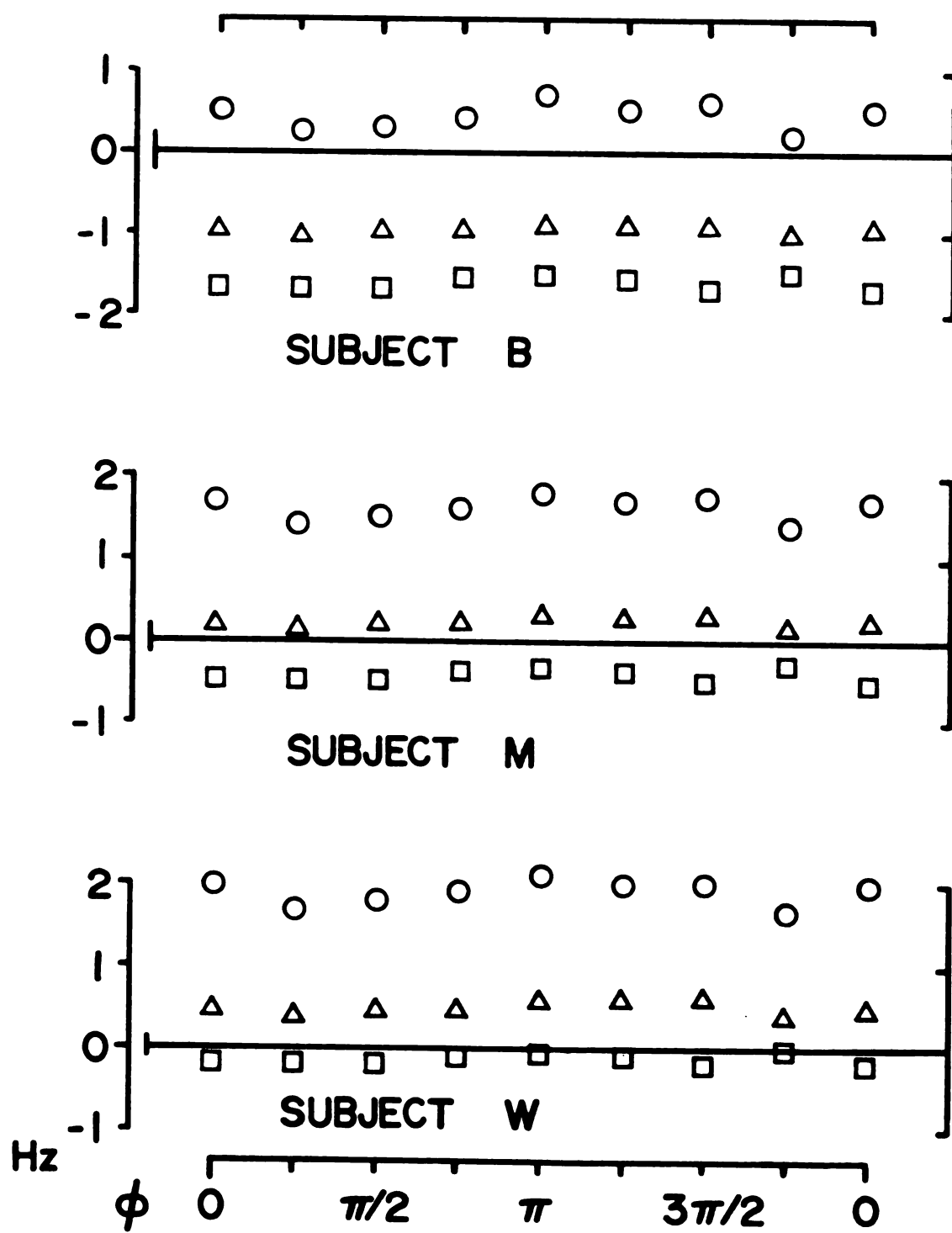


Figure 14. Amplitude of 12 Hz component relative to thresholds.

Table 4. FM detection levels in Hertz for 4 and 12 Hz modulation rates.

		Subject	Mean	Std. Error
Modulation rate				
4 Hz		B	1.846	.165
		M	1.831	.094
		W	.788	.021
		MEAN	1.488	
<hr/>				
12 Hz		B	2.416	.193
		M	1.244	.184
		W	.926	.165
		MEAN	1.529	

models. If the discrimination results cannot be predicted spectral type models can be rejected.

From Figure 14 we see, then, that subject W should be able to perform the discrimination task with virtually no errors for $\Delta f_g = +5$ and $+10$ Hz. He should perform at the chance level for $\Delta f_g = +2.5$ Hz. Subject M should perform with no errors at $\Delta f_g = +10$ Hz, only slightly better than chance at $\Delta f_g = +5$ Hz and at chance for $\Delta f_g = +2.5$ Hz. Subject B should perform only slightly better than chance for $\Delta f_g = +10$ Hz and at chance for both $\Delta f_g = +5$ Hz and $\Delta f_g = +2.5$ Hz. For subject W (Figure 12.) the order of the results matches the predictions but the fine details are unaccounted for. For $\Delta f_g = +2.5$ Hz, the discrimination data shows essentially chance performance. It might be noted that the extremely small variations in the level of the third harmonic are paralleled in the discrimination data. This seems to be a fortuitous effect attributable solely to chance. The level of performance is so low that most variations must be due totally to random error. At $\Delta f_g = +5$ Hz Subject W did not perform without error as expected. The variations in performance were large enough to rule out the possibility of being completely due to random error and were unmatched by the fine structure of the actual levels of the third harmonic. At $\Delta f_g = +10$ Hz the performance was nearly error-less with a large decrease as ϕ approaches π . It might be argued that this is due to changes in the size of the third harmonic, but at $\phi = \pi$ the third harmonic is at its largest and certainly this should not predict a decrease in performance.

Subject M performed much like predicted (see Figure 11). At $\Delta f_g = +2.5$ Hz the data are at chance level. At $\Delta f_g = +5$ Hz the data are much more variable and just slightly above chance. At $\Delta f_g = +10$ Hz

performance is virtually errorless.

Subject B (Figure 10) also performed nearly as expected. For $\Delta f_s = \pm 2.5$ and ± 5 Hz the data are very close to chance level. For $\Delta f_s = \pm 5$, however, at ϕ approaching 0 there seemed to be a consistent increase in performance. This is difficult to explain with the third harmonic so far below threshold for these stimuli. At $\Delta f_s = \pm 10$ Hz there is again shown an unexplainable drop in performance when the third harmonic is at its highest amplitude.

Clearly, at least for subjects B and W there is more being processed than the amplitude of the third harmonic. The temporal structure of the modulation waveform must be affecting the discrimination process. the spectral-type models can be rejected for this task.

Experiment IV

As described in Experiments I, II, and III there is a significant dependance of modulation perception on the phase relationship between the two components of the complex modulation waveform. Experiment I showed an asymmetric phase effect about $\phi = \pi$. Experiments II and III showed a phase effect that is symmetric. These effects cannot be explained by spectral-type models. Experiment IV is an attempt to test a specific temporal model, the 1-pole Lag-Processor model.

An informal visual analysis of the waveforms (Figure 2) might lead one to guess that if a phase effect is going to be found it would be symmetric about either $\phi = \pi$ or $\phi = 0$, the "extremes" of the waveforms. The waveforms at intermediate values of ϕ seem to form symmetric progressions from one extreme to the other. As a consequence of changing ϕ the relationship between the peaks of the third harmonic component and the peaks of the fundamental component of the complex modulation

change. The variability in the location of the irregularity in the summed modulation waveform is due to that changing relationship. As ϕ approaches π this irregularity seems to move further away from the "global maximum", occurring progressively earlier. For ϕ greater than π , the irregularity is on the trailing side of the global maximum and seems to move closer to the peak as ϕ increases. The actual value for the global maximum also changes with ϕ , but this change is symmetric about $\phi = \pi$ and would not be expected to have an asymmetric effect.

The Lag-Processor model, that Hartmann (1978) discarded, does predict an asymmetric phase effect and so provides a place to begin further investigation. By this model, the ear takes a weighted average of all frequency deviations occurring during a moving "time window" of constant "length". The frequency deviation occurring at the leading edge of that window, (i.e. the present instant), is given a maximum weighting. All deviations occurring earlier in time are given progressively less weight, an exponentially decaying function of time. All deviations occurring at the same relative time prior to the leading edge of the moving window are given the same weight. The sequence with the largest maximum in its moving average would be perceived as having the widest frequency deviation.

This model predicts that a sequence of frequency deviations with a small peak preceding a large peak will yield a higher maximum in its moving average than a sequence of frequency deviations where the order of those peaks is reversed. The result depends, however, on the length of the time window. When the pattern of weights decreases to zero too quickly only one of the peaks may be captured within the window. In this case, since the size of the large peak is the same for both sequences, the maximum moving average would also be the same. As the

length of the time constant increases to include more of the second peak the difference between maximum moving averages will increase. The difference will increase to some maximum and then begin to decrease as the time window continues to lengthen. The difference will decrease as the two peaks become more similar due to the time constant increase. Figure 15 shows the relationship between an exponentially decaying time window and the complex waveform for $\phi = \pi/4$ and $\phi = 7\pi/4$ for $f_m = 2$ Hz, 4 Hz, and 8 Hz. The time constant is based on data collected by Kay and Matthews (1972).

Kay and Matthews (1972) collected data showing thresholds for detection of FM at various modulation rates. The response curve matched what might be expected if FM was processed through some type of low-pass filter. The thresholds for slow modulation rates were quite low. Those thresholds increased as the modulation rate increased. Those data can then be used to determine the cut-off frequency of the low-pass filter. From that cut-off frequency the time constant of the 1-pole Lag-Processor can be determined. We estimate that time constant to be 40 msec.

The time window in the figures is drawn to give maximum weight to the leading peak. For $\phi = \pi/4$ this is near the point at which the maximum average is achieved. Because of the increasing time separation between the peaks as f_m goes from 8 Hz to 2 Hz, the preceeding peak is given progressively less weight. By informal visual analysis of these graphs it appears that the largest differences between the maximum moving averages will be produced by $f_m = 8$ Hz. At $f_m = 2$ Hz the computed average is obtained almost entirely from a single peak. The mechanism will obtain its maximum output when the leading edge is at or near the maximum peak for both waveforms. This should yield a relatively small

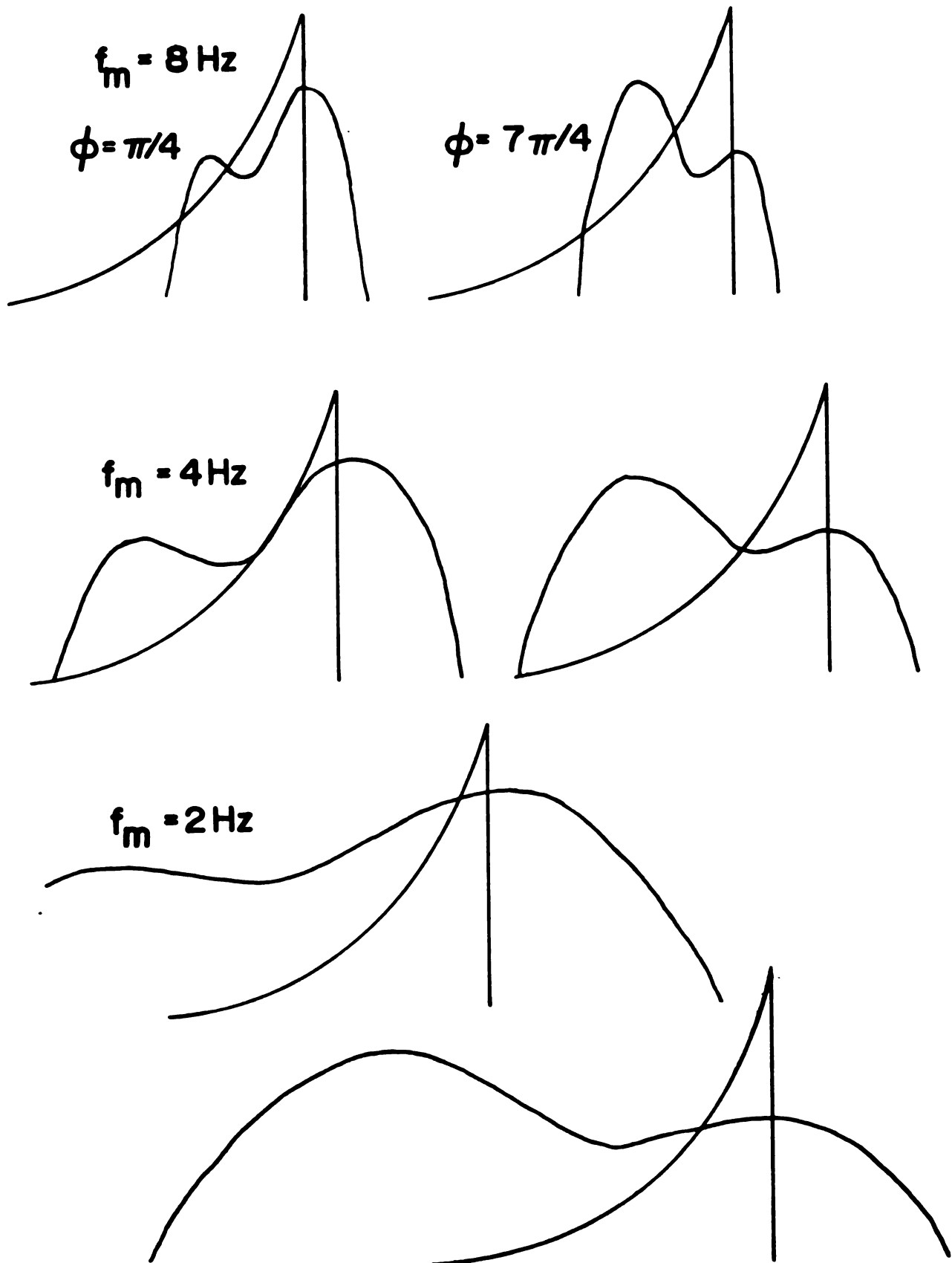


Figure 15. Exponential windows superimposed on modulation waveforms.

difference in perceived width between the otherwise symmetric waveform pair.

The Q_e -value provides a measure of relative width. The Q_e -value is computed as a ratio of actual frequency deviations at PSE for width. A smaller Q_e -value, therefore, indicates that a particular modulation waveform is perceived as relatively wider than a sine based FM with an equal actual frequency excursion. Since all complex FM waveforms were compared to equal-width-sine-waveform modulation, Q_e -values may be compared across phase angles. The asymmetry of Experiment I can, therefore, be roughly explained by this application of the Lag-Processor model.

This relationship between the parameters of the peaks and the length of the time constant may be investigated by manipulating any of the following variables: the length of the time constant, the time separation between the peaks, the size of the peaks, the order of the peaks, and the pattern of the weights within time window. Changing any one or any combination of these variables will affect the size of the moving average derived from the time window. In the context of the human auditory system, however, we assume that the time constant and the pattern of weights within the time window are fixed.

Experiment I varied the order and relative sizes of the peaks of the waveform. To further test this model Experiment IV varies the time separation between the peaks of the waveform relative to the fixed length of the time window to the auditory system postulated by the model. This is accomplished by varying the modulation rate, f_m .

This provides a test of a temporal-type model. If the time constant derived from the Kay and Matthews (1972) data is correct for these subjects, the Lag-Processor model predicts that the differences

between the Q_e -values obtained at each waveform group should be relatively larger for $f_m = 8$ Hz than for $f_m = 2$ Hz. Q_e -values for $\phi > \pi$ should continue to be larger than for $\phi < \pi$.

Because of the decaying weight pattern it is the case that the effect of the smaller peak on the moving average will decrease rapidly as the small peak decreases in size. This leads to the additional prediction that the difference between the Q_e -values at $\phi = \pi/4$ and $\phi = 7\pi/4$ should be larger than for $\phi = 3\pi/4$ and $5\pi/4$, at $f_m = 8$ Hz. This effect is not expected for $f_m = 2$ Hz because the time constant should be too short relative to the structure of the modulation waveforms. This is the same reason for the predicted differences between 8 and 2 Hz FM above.

Method. The method for this experiment is identical to Experiment I with the following parameter changes:

- 1.) Two values of f_m were used, $f_m = 2$ Hz and 8 Hz.
- 2.) Only one value of Δf_s was used, $\Delta f_s = \pm 5$ Hz.
- 3.) Only 6 different phase angles were used:

$$\phi = \pi/4, \pi/2, 3\pi/2, 5\pi/4, 3\pi/2, \text{ and } 7\pi/4.$$

and 4.) Only subjects B and M participated.

Results - Discussion. The measured Q_e -values are shown in Figure 16.

It is quite obvious that the expected asymmetries did not occur. Table 5 shows the 95% confidence intervals for the difference between the data at large values of ϕ versus small ϕ values. The $\Delta f_s = \pm 5$ Hz data of Experiment I are also included for comparison. Neither subject showed a significant asymmetry for $f_m = 8$ Hz, while Subject B showed an asymmetry for $f_m = 2$ Hz. The predictions were not matched by the

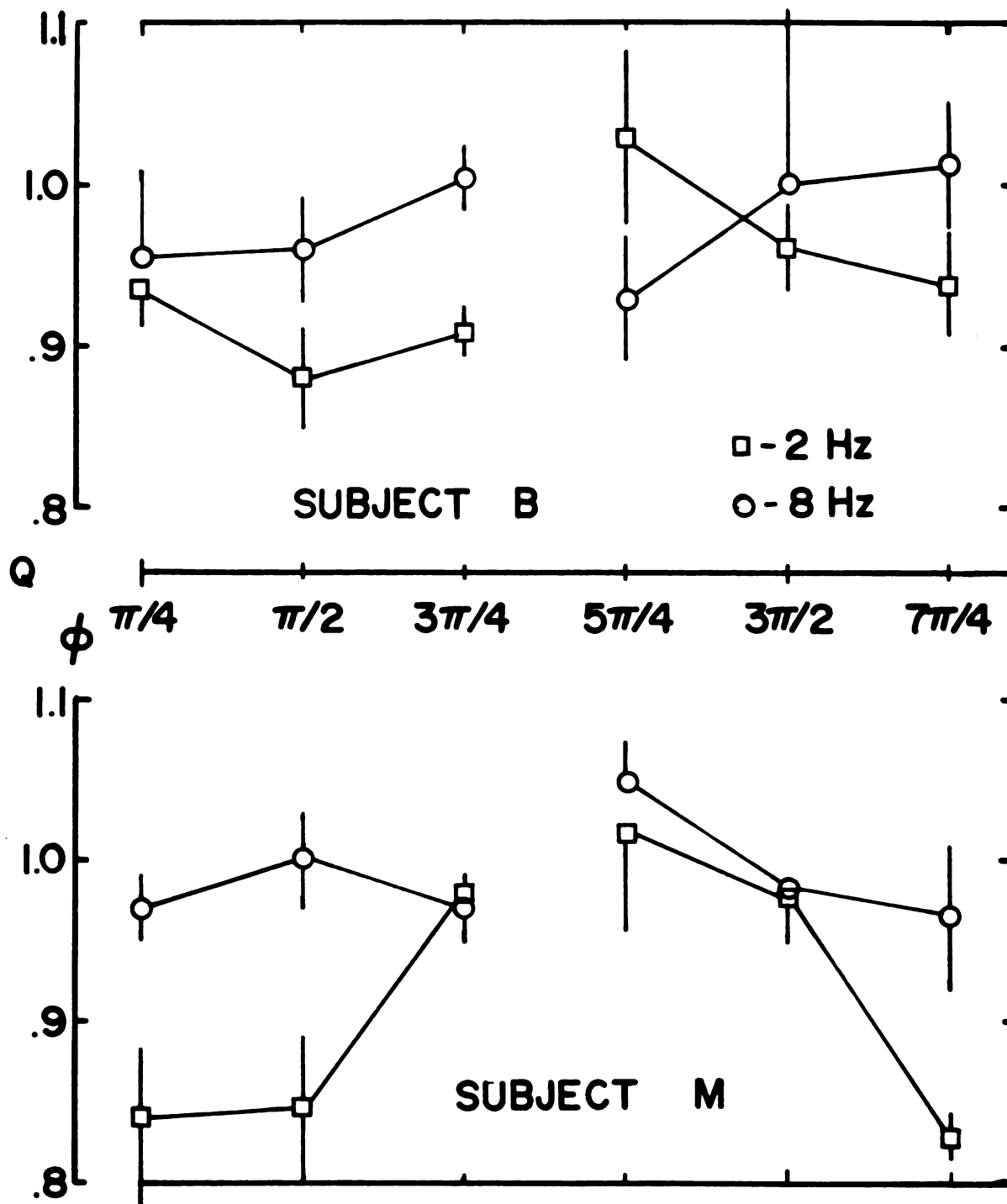


Figure 16. Q_0 -values. Experiment IV data.

Table 5. 95% confidence intervals for differences in Q_e -values
for $\phi < \pi$ versus $\phi > \pi$.

Subject	Modulation rate:		
	2 Hz	4 Hz	8 Hz
B	$.021 < \delta < .115$	$-.010 < \delta < .060$	$-.071 < \delta < .073$
M	$-.035 < \delta < .139$	$.037 < \delta < .113$	$-.022 < \delta < .062$

$$\delta = Q_{e(\phi > \pi)} - Q_{e(\phi < \pi)}$$

data. It must be noted, however, that those predictions depended heavily upon the time constant obtained from the Kay and Matthews (1972) data. It might be the case that the individual variations in that time constant are so great as to yield the asymmetries that were seen. For Subject B, then, the time constant might be very long. This would shift the asymmetry effect down to a much lower modulation rate. For the higher modulation rates no asymmetries would appear because the entire modulation waveform may be too close to the leading edge of the time window. For Subject M the same might also be true, but the effect centered around a slightly shorter time constant. If this interpretation is correct this procedure provides a way of discovering an individual's time constant. This conclusion must be verified by some other method of estimating individual time constants. If this cannot be verified this data serves as evidence against this specific temporal model. Some alternative must be devised.

The original time constant was generated from FM detection threshold data. A small portion of the necessary data has already been collected for these subjects in Experiment III. With only 2 modulation rate detection levels collected for each of the two subjects no quantitative measurements can be given. It is possible to work "backwards" from the data of Experiment IV and predict that Subject B should have a higher threshold for $f_m = 4$ Hz than Subject M and a much higher threshold at $f_m = 12$ Hz. This is verified by the data. Subject M should not have a lower threshold at $f_m = 12$ Hz than at $f_m = 4$ Hz, however.

The second prediction must now be modified in light of the above discussion. If the time constant variations cannot be further verified the prediction of smaller differences at $\phi = 4\pi/4$ and $5\pi/4$ has no basis. If the model were found to hold up these predictions would also be

adjusted by the individual subjects' time constant. For Subjects B and M this means that more than one modulation cycle will be captured within the window for $f_m = 8$ Hz. Without a specific measure of the time constant it is impossible to predict the differences across phase angles.

It might be noted from Figure 16 that at $f_m = 2$ Hz for Subject B the difference between Q_e -values at $\phi = \pi/4$ and $7\pi/4$ is small compared to the difference at $\phi = 3\pi/4$ and $5\pi/4$. At $f_m = 8$ Hz Q_e for $\phi = 7\pi/4$ may be larger than at $\phi = \pi/4$, but Q_e at $\phi = 5\pi/4$ is actually considerably less than for $\phi = 3\pi/4$. These differences are not predicted at all by the original time constant. For Subject M the differences between Q_e -values at $\phi = \pi/4$ and $7\pi/4$ are quite small while at $\phi = 3\pi/4$ and $5\pi/4$ the differences are relatively large. This is the case for both modulation rates. This is also not according to the predictions based on the original time constant.

The 1-pole Lag-Processor model appears to be inadequate to explain all of the structure in the data. More detailed measurements of individual time constants may provide better parameters to help this model, but those adjustments will probably not be enough.

General Discussion and Conclusions. This series of experiments has been an attempt to discover the character of the processing underlying our perception of frequency modulated auditory signals. Given the early findings that the human auditory system does not follow the frequency deviations strictly, some process other than simple frequency following must be utilized.

There are two approaches to modelling this perceptual process. The first is labeled the spectral view in which some combination rule for the independently detected Fourier components of the modulation waveform

determines the perceived modulation width. This takes place without processing information regarding the phase relationships between the components. The other view is a catch - all category including all models that do not rely upon the Fourier analysis of the waveforms. This group of models makes use of the actual time-varying microstructure of the modulation waveforms and is called the temporal view.

Recent work has outlined a technique for obtaining a useable measure of perceived modulation width. Use of this method has accompanied tests, primarily of spectral models for explaining the acquired data. A major problem has been the use of spectrally complex modulation waveforms. The presence of many spectral components makes interpretation of results very difficult. This study has avoided that problem by using only spectrally simple modulation waveform.

Experiment I was a measurement of the effect of phase angle and overall amplitude on the perceived width of complex waveform FM. A significant asymmetric phase effect was found providing evidence against the spectral-type model. This finding was further investigated in Experiment IV. The second finding was a dependence of relative width on overall excursion of the FM waveform. This could be interpreted both spectrally and temporally. Experiment II was to investigate this effect further.

Experiment II measured the discriminability of the various complex FM waveforms from sine-waveform FM. The waveforms used covered the same parameters as were used in Experiment I. It was found that discrimination performance improved as overall modulation excursion increased, but that it varied with the phase relationship between the complex waveform components. These effects indicated that the same process mediating FM width perception also played a part in FM waveform discrimination.

It was unclear, however, whether or not the phase effect was due to the variations in the size of the third harmonic necessary to equalize the FM signals for equal width. It was determined that the relevant factor would be the deviation of the third harmonic amplitude relative to the subjects' threshold for 12 Hz FM detection.

Experiment III was the measurement of the individual subjects' FM threshold for the third harmonic modulation rate. Application of this data to the data of Experiment II showed that the deviations from threshold of the third harmonic component explained the large, overall size effect, but were diametrically opposed to the changes in discrimination performance associated with phase angle. This led to the further rejection of the spectral-type model of FM width perception and waveform identification.

Experiment IV was a test of a 1-pole Lag-Processor temporal model for explaining the asymmetry of the phase effect in Experiment I. Using a time constant drawn from data of Kay and Matthews (1972) as a parameter the model failed to predict asymmetries in width perception for 2 values of modulation rate. Alterations of the time constant parameter would correctly explain the asymmetries, but could not deal with the fine structure of the data. This particular temporal model was rejected.

In summary, spectral-type models of FM width perception and waveform discrimination were completely rejected. The Lag-Processor, a specific temporal-type model was found inadequate to explain the asymmetric phase effect in the width perception data.

LIST OF REFERENCES

LIST OF REFERENCES

- Blakemore, C.B. & Campbell, F.W. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. Journal of Physiology, 1969, 203, 237-260.
- Van Den Brink, G. Experiments on binaural diplacusis and tone perception. In R. Plomp & G. Smoorenburg (Eds.), Frequency analysis and periodicity detection in hearing. Leiden: A. W. Sijthoff, 1971.
- Van Den Brink, G. The influence of fatigue upon the pitch of pure tones and complex sounds. Presented at Symposium on Hearing Theory, Eindhoven, Holland, 1972.
- Deutsch, D. Octave generalization and tone recognition. Perception and Psychophysics, 1972, 11, 411-412.
- Divenyi, P.L. & Hirsh, I.J. Pitch changes in trills and vibrato. Journal of the Acoustical Society of America, 1972, 51, A138.
- Dowling, W.J. The perception of interleaved melodies. Cognitive Psychology, 1973, 5, 322-337.
- Graham, N. & Nachmias, J. Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channels models. Vision Research, 1971, 11, 251-259.
- Hartmann, W.M. Five experiments on frequency modulation width perception. Journal of the Acoustical Society of America, 1977, 61, S50.
- Hartmann, W.M. Perception of frequency modulation width. Unpublished manuscript, Michigan State University, 1978.
- Hartmann, W.M. & Long, K.A. Time dependence of pitch perception-vibrato experiment. Journal of the Acoustical Society of America, 1976, 61, S50.
- Kay, R.H. & Matthews, D.R. On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones. Journal of Physiology, 1972, 225, 657-677.
- Klein, M.A., Gable, G.A., Edmends, D.L., Eicher, D.A., & Hartmann, W.M. Microcomputer control for psychoacoustic experiments. Journal of the Acoustical Society of America, 1978, 63, S63.

- Kock, W.E. Certain subjective phenomena accompanying a frequency vibrato. Journal of the Acoustical Society of America, 1936, 8, 23-25.
- Lewis, D. Cowan, M., & Fairbanks, G. Pitch and frequency modulation. Journal of Experimental Psychology, 1940, 27, 23-36.
- Nabelek, I.V., Nabelek, A.D., & Hirsh, I.J. Pitch of sound bursts with continuous or discontinuous change of frequency. Journal of the Acoustical Society of America, 1973, 53, 1305-1312.
- Seashore, C.E. (Ed.) The vibrato. University of Iowa Studies in the Psychology of Music (Vol. 1). Iowa City: The University Press, 1932.
- Seashore, C.E. The psychology of the vibrato. University of Iowa Studies in the Psychology of Music (Vol. 3). Iowa City: The University Press, 1935.
- Seashore, H.G. The hearing of pitch and intensity in vibrato. In C.E. Seashore (Ed.) University of Iowa Studies in the Psychology of Music (Vol. 1). Iowa City: The University Press, 1932.
- Sundberg, J. Vibrato and vowel identification. Archives of Acoustics, 1977, 2, 257-266.
- Sundberg, J. Effects of the vibrato and the 'singing formant' on pitch. Proceedings of Musicologica Slovaca, VI, Bratislava, 1978.
- Zwicker, E. Die grenzen der horbarkeit der amplitudenmodulation und der frequenzmodulation eines tones. Acustica, 1952, 2, AB125-AB133.
- Zwicker, E. Temporal effects in psychoacoustical excitation. In A.R. Moller (Ed.) Basic Mechanisms in Hearing. New York: Academic Press, 1973.