INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

77-17,990

ERDREICH, John, 1943-
SUBJECTIVE TONES.

The University of Michigan, Ph.D., 1977
Psychology, psychobiology

Xerox University Microfilms, Ann Arbor, Michigan 48106

SUBJECTIVE TONES

by

John Erdreich

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physiological Acoustics)
in The University of Michigan
1977

Doctoral Committee:

Associate Professor T. Dean Clack, Chairman
Associate Professor David J. Anderson
Professor Merle Lawrence
Professor William R. Uttal

DEDICATION

In appreciation of my family who made this dissertation possible and of my teachers who made it worthwhile.

ACKNOWLEDGEMENTS

This study was supported, in part, by funds from the following sources:

Public Health Service Grants NB07352 and NF05785

The University of Michigan Medical School Fund for Computing

Horace H. Rackham School of Graduates Studies Dissertation Grant.

Special thanks are due Mr. James Bruce, Kresge Hearing Research Institute, Graphic Arts Staff for the execution of the drawings in this dissertation and to Mrs. Rita Walker for the typing of the manuscript.

TABLE OF CONTENTS

Page

DEDICATION.....ii

ACKNOWLEDGEMENTS.....iii

LIST OF TABLES.....vii

LIST OF FIGURES.....viii-ix

CHAPTER ONE INTRODUCTION.....1

I. Subjective Tones 1

A. Periodicity Pitch.....3

B. Beat-Tone Theory 3

C. Amplitude Distortion.....5

II. The Basis of the Controversy.....6

III. Related Issues.....9

A. Monaural Phase Effects.....9

B. Subjective Tone Masking.....10

IV. Three Experiments.....11

CHAPTER TWO AURAL HARMONIC ESTIMATION.....13

I. Methods.....13

A. Subjects.....13

B. Stimulus Production.....14

C. Aural Harmonic Estimation.....18

D. Combination Tone Level Prediction.....20

II. Measurements of an Harmonic Maskee in the Presence of a Fundamental.....21

A. Thresholds.....21

B. Distortion Products.....23

C. Predicted Combination Tone Levels.....26

Page

III. DISCUSSION.....28

A. The Variation of Masked Threshold with Phase.....28

B. Growth of the Aural Harmonic.....28

C. Masking at the Harmonic.....29

D. Masker Level Dependence of Parameters of the Nonlinearity.....30

CHAPTER THREE ESTIMATION OF COMBINATION TONES WITH THE TONE-ON-TONE MASKING PARADIGM.....31

I. Methods.....32

II. Masked Threshold Measurement Above the Frequency of the Primaries.....33

A. Monaural Phase Effects.....33

B. Masked Thresholds.....33

C. Summation Tone Estimates.....36

D. Effects of Frequency Separation.....43

III. Estimates of f_h-f_1 with Tone-on-Tone Masking.....44

IV. Discussion.....47

CHAPTER FOUR MEASUREMENT OF THE AUDIBLE DIFFERENCE TONE $2f_1-f_h$.....55

I. Methods.....56

A. Cancellation Paradigm.....56

B. Loudness Balance Paradigm.....60

II. Level of the Difference Tone $2f_1-f_h$....63

Page

III. Discussion.....67

CHAPTER FIVE THREE MODELS OF AUDITORY NONLINEARITY...70

I. Helmholtz' Theory.....70

II. Fletcher's Model.....72

III. Essential Nonlinearity.....74

IV. Conclusions.....77

CHAPTER SIX SUMMARY.....79

APPENDIX ESTIMATE ACCEPTABILITY (ACCEPTANCE CRITERION RATIO).....81

BIBLIOGRAPHY.....84

LIST OF TABLES

Table		Page
2.1	Stimulus Distortion Levels	16
2.2	Aural Harmonic Estimates and Power Series Coefficients	27

LIST OF FIGURES

Figure		Page
2.1	Equipment Configuration	15
2.2	Threshold Shifts at the Harmonic	22
2.3	Masked Thresholds at the Harmonics	24
2.4	Aural Harmonic Estimates	25
3.1	Threshold Shifts for Summation Tones (f_h/f_1 = 1.2)	34
3.2	Threshold Shifts for Combination Tones (f_h/f_1 = 1.6)	35
3.3	Masked Thresholds for Summation Tones (f_h/f_1 = 1.2)	37
3.4	Combination Tone Masked Thresholds (f_h/f_1 = 1.6)	38
3.5	First Order Summation Tone Estimates (f_h/f_1 = 1.2)	39
3.6	Second Order Summation Tone Estimates (f_h/f_1 = 1.2)	41
3.7	Combination Tone Estimates (f_h/f_1 = 1.6)	42
3.8	Frequency Effects on Distortion Product Level	45
3.9	Measures of (f_h-f_1)	46
4.1	Recording Attenuator Tracing	58
4.2	Representation of Tones in Cancellation	59
4.3	Stimulus Sequences	62
4.4	Equal Loudness Judgement Controls	64

Figure Page

4.5 Cancellation Measures of $2f_1-f_h$.....65

4.6 Uncorrected Cancellation Measures of $2f_1-f_h$.....66

4.7 Loudness Balance Estimates of $2f_1-f_h$.....68

CHAPTER ONE

INTRODUCTION

Sensory processes, in addition to providing information about the environment, may introduce spurious information. These distortions which result from imperfections in the sensory system are important because of the insights into the mechanisms of sensation and perception that may be gained by analysis of sensory experience having no physical stimulus correspondence. In audition, subjective tones are examples of sensory misrepresentation.

In this investigation we studied the existence and nature of subjective tones and have attempted first to demonstrate three subcategories of subjective combination tones: aural harmonics, summation tones, and difference tones. We then describe the subcategories and compare the characteristics of each class of subjective tone to determine whether the behavior of tones in each subcategory might be described by a common model or whether separate nonlinear mechanisms are needed. The third consideration is the comparison of three models of amplitude distortion which explain the origin of these subjective tones: the nonlinear theory of Helmholtz (1854), the power-series nonlinearity of Fletcher (1929) and the essential nonlinearity theory (Goldstein, 1967).

I. Subjective Tones

Although the existence of low pitch subjective tones produced by two higher pitch notes has been acknowledged for

nearly two centuries (cf. Jones, 1935), their mode of generation has been uncertain and is still the subject of controversy. These subjective tones are heard as if they are acoustic signals but are characterized by the lack of a physical component of corresponding frequency in the stimulus.

Two tones will combine to yield a series of subjective tones. The frequencies of the individual members (f_{DP}) of the combination tone series may be represented in terms of the higher (f_h) and lower (f_1) primary frequencies as

$$f_{DP} = mf_1 \pm kf_h \qquad m, k = 0,1,2,\ldots, n.$$

The order of a combination tone may then be represented as $(m + k - 1)$ (Hällström, 1832).

As an example, the subjective tones under consideration in this study are the first order harmonic of 1000 Hz, $2f_1$ = 2000 Hz and the first order summation tone f_h+f_1 = 2200 Hz (f_1 = 1000 Hz and f_h = 1200 Hz). Second order tones investigated were the third harmonic of f_1, $3f_1$ = 3000 Hz, the summation tone $2f_1+f_h$ = 3200 for f_1 = 1000 Hz and f_h = 1200 Hz, $2f_1+f_h$ = 3600 Hz with f_h increased to 1600 Hz, $2f_h+f_1$ = 4200 Hz with the same primary tone frequencies and the second order difference tone $2f_h-f_1$ = 2200 Hz again with f_1 = 1000 Hz and f_h = 1600 Hz.

There are at least three explanations for these auditory phenomena: periodicity pitch production, beat-tone theory, and amplitude distortion. The first two explain only difference tones while the last one also explains harmonics and summation tones.

A. Periodicity Pitch

A complete description of the phenomenon of periodicity pitch may be found in a review by Small (1970). This subjective pitch has been related to the time structure of the acoustic stimulus.

The apparent pitch of a complex acoustic signal is related to that frequency the period of which corresponds to the largest common period of the signal. Thus, sensation of pitch may be elicited by sequences of acoustic pulses, by periodically interrupted noise, or by modulated pure tones. In each case, there is an absence of energy in the stimulus at frequencies of the observed pitch.

Although the pitch of the first order difference tone could be explained in the context of periodicity pitch theory, it cannot explain higher order difference tones, aural harmonics, or summation tones. For this reason it will not be considered in explanation of the subjective tones being studied.

B. Beat-Tone Theory

A theoretical explanation of subjective tone production embracing all difference tones was postulated by Thomas Young (1800). Based on the impression that the beats generated by two tones gradually changed into a difference tone as the frequency separation of the two tones increased, Young proposed both beats and difference tones as evidence of a linear superposition (addition) of two tones:

> "vibrations constituting different sounds should be able to cross each other...without disturbing each other's progress; but this can be no otherwise affected than by each particle's partaking of both motions. If this assertion stood in need of any proof, it might be amply furnished by the phenomena of beats, and of the grave harmonics observed by Romieu and Tartini...."

This directly then is the foundation upon which beat-tone theory developed. Young proceeds:

> "The greater the difference in the pitch of two sounds, the more rapid the beats, till at last, ...they communicate the idea of a continued sound; and this is the fundamental harmonic described by Tartini."

Difference tones, he explains, are the consequence of stimulus envelope variations becoming too rapid to be separately perceived. Young's theory relates the waveform over many cycles to the perceived pitch, in contrast with the stochastic concerns of periodicity pitch theories.

In addition to the first difference tone, Young reported the occasional observation of a secondary subjective tone. The origin of these and other difference tones were explained by Hällström (1832) with a reinterpretation and modification of Young's beat-tone theory. Higher order difference tones, he explained, were produced by interaction of lower order

difference tones and primaries. Thus, the interaction of f_h-f_l with f_l generated $2f_l$-f_h. Higher order difference tones were produced by similar interactions.

Implicit in this interaction mechanism then is the physical realization of the difference tone. In the strict sense, however, the interaction of the two beating tones according to Young is additive and therefore linear (i.e., $Y = A \sin f_l + B \sin f_h$). The mechanism cannot produce a physical subjective tone. There should be no "beat-tone" with which a primary can interact to produce a higher-order difference tone. Furthermore, the theory cannot explain the existence of harmonics or of summation tones.

The primary contribution of Young and Hällström was not a viable explanation of the subjective tones under consideration here but rather they were the first to formalize the idea that the subjective tones were the result of a physical interaction of two tones. Here too was the first suggestion that the subjective tones themselves may interact to produce a new generation of subjective pitch. Beat-tone theory, however, does not explain the subjective pitches we are studying.

C. Amplitude Distortion

In his discussion of combination tones, Helmholtz (1855) postulated that the difference tones reported by Sorge and Tartini a century earlier were the result of a nonlinear amplitude characteristic of the middle ear. Therefore, in addition to distortion products (DP) at the difference freq-

uencies

$$f_{DP} = nf_1 - mf_h \qquad n,m = 1,2,3,\ldots,k$$

he reasoned that two other classes of subjective tone must exist. These were aural harmonics and summation tones described respectively as

$$f_{DP} = nf_1 \text{ or } mf_h \text{ and}$$

$$f_{DP} = nf_1 + mf_h.$$

Although Helmholtz claims to have heard these tones, he was the only one to have done so.

His theory of nonlinearity assumed that in response to a displacement a restoring force was produced proportional to the square of the displacement. This resulted in subjective tone amplitudes related to both the frequencies of the primaries and the amplitude of the stimulus.

In 1929, Fletcher simplified Helmholtz' model. From his modification it may be predicted that equal-order combination tones (e.g. f_h-f_1 and f_h+f_1; $2f_1 \pm f_h$ and $2f_h+f_1$) are equal in amplitude and furthermore, for equal stimuli to the nonlinearity, the levels of subjective tones of the same order are frequency independent. Although amplitude nonlinearity was useful in explaining all three subcategories of subjective tone, the mechanism of subjective tone generation was still not clearly established.

More contemporary studies would both support and question the concepts of nonlinearity proposed above.

II. The Basis of the Controversy

Psychoacoustic studies commencing in the 1920's tended

to confirm the existence of harmonics and summation tones although the summation tones were never actually measured. Wegel and Lane (1924) observed beating at harmonic frequencies of a masker which they attributed to interaction of a harmonic of the masker and the maskee. They calculated that the amplitude of the harmonic increased faster than the increase of masker level. Chapin and Firestone (1934) demonstrated constructive and destructive interference between a subjective harmonic and a pure-tone maskee at the same frequency. In 1972 estimates of aural harmonic amplitudes were reported by Clack, et. al. These harmonic levels exhibit, within the range of accelerated growth, increasing amplitudes of approximately 2dB/dB for the second harmonic and 3dB/dB for the third harmonic (see Clack, et. al., 1972; Clack, 1975). Although these measurements tended to support the power-series hypothesis, two problems still existed.

One was simply that summation tones had still not been demonstrated. The other arose from measurements of the audible subjective tone $2f_1-f_h$. Zwicker (1955) made measurements of the subjective tones f_h-f_1 and $2f_1-f_h$. Although the growth of f_h-f_1 appeared to be in accord with power-series predictions, the difference tone $2f_1-f_h$ increased at 1dB/dB rather than 3dB/dB as he predicted based on a third-order nonlinearity. These results were confirmed and extended by Goldstein (1967).

To explain the growth of $2f_1-f_h$ and the inaudibility (or possible nonexistence) of the summation tones and

harmonics, Goldstein developed a new theory of auditory nonlinearity. His essential nonlinearity hypothesis consists of the following points: (1) the nonlinear nature of the auditory response was an amplitude independent essential nonlinearity. This form of nonlinearity contains a factor (PK) such that the contribution of higher ordered term in the series

$$X = a_0 + a_1 p + a_2 \frac{p^2}{PK} + a_3 \frac{p^3}{PK^2} + \ldots + a_n \frac{p^n}{PK^{n-1}}$$

are normalized with respect to peak amplitude (PK is the peak pressure of the stimulus and "p" is the stimulus sound pressure level); (2) distortion products higher in frequency than the primaries are reduced by a filter following the nonlinearity and are not perceived.

A nonlinear mechanism describable by Goldstein's equation produces a distortion product at $2f_1 - f_h$ equal to $p_1^2/(p_1 p_2)^2$, in agreement with his data and that of Zwicker.

Although both hypotheses lead to different predicted subjective tone characteristics, both power-series and essential nonlinearity hypotheses have similar aspects. Each proposed a mathematical principle that governs the behavior of subjective tones. Each is based on the assumption that sum and difference tones arise by some nonlinear interaction of two or more stimuli. Each has a common dependence of distortion product amplitude on signal level. Additionally, by their definition, each of the alternative hypotheses is frequency independent with respect to the input to the nonlinearity.

The differences between the predictions of subjective tone amplitudes are that essential nonlinearity leads to a 1dB/dB growth of a difference tone of any order and ineffective sum tones and harmonics, whereas power-series nonlinearity leads to subjective tone growth proportional to the order plus one of a particular subjective tone.

The question of paramount concern is: must two mechanisms be postulated to explain the observed subjective tone behavior? If summation tones are found to be non-existent this would be evidence in favor of Goldstein's hypothesis of auditory nonlinearity. If summation tones exist, however, the previously reported measures of difference tone characteristics must be carefully scrutinized, validated, and compared with the characteristics of other subjective tones. Any basis on which these estimates could be in error would be a strong point from which to develop support of the power-series nonlinearity.

III. Related Issues

A. Monaural Phase Effects

Interaction of a maskee tone at a distortion product frequency and a fundamental frequency primary tone or tones which produces a sinusoidal change in threshold for the maskee (the monaural phase effect) has been traditionally considered evidence of the existence of a subjective tone (Chapin and Firestone, 1934; Wegel and Lane, 1924; Lawrence and Yantis, 1956; Bekesy, 1960; Clack, 1967, 1968; Clack, et. al., 1972; Erdreich and

Clack, 1972). The interaction is thought to be the result of the vector summation of the stimulation resulting from the maskee tone with that resulting from the distortion product (Lewis and Larsen, 1937) and has been the foundation on which most distortion product measurements have been based. Alternative explanations of the monaural phase effect have been formulated considering changes in the waveform envelope as individual stimulus components interact (Plomp, 1967), changes in stimulus fine time structure (Zwicker, 1976),and detection of changes in the relative phases of the components (de Boer, 1961).

In the discussion of the summation tone measurements tone-on-tone masking data from this and other studies will be examined to determine which of the alternative explanations of the monaural phase effect are viable and whether the existence of the monaural phase effect is evidence of aural summation tones.

B. Subjective Tone Masking

Accepting the validity of the subjective tone estimates, another question may be raised. The masking of high frequency tones by a low frequency maskee is known to increase at approximately 2dB/dB (Ehmer, 1959). How are we to determine then that the 2dB/dB growth of second order harmonics and other quadratic distortion products are the result of a nonlinear process independent of masking? One approach is to compare masking and growth

of the same subjective tone. If the masking and growth are identical, then further tests need to be devised to answer the question. If, however, they differ or if the masking of a second-order tone such as $2f_l \pm f_h$ is also 2dB/dB, but the growth is 3dB/dB, then masking and amplitude distortion are independent.

IV. Three Experiments

In the first experiment, the masked threshold for a tone corresponding to either the second or third harmonic of a masker is measured as a function of phase shift of the maskee. From these measurements estimates of the level of the second and third aural harmonic are determined and the parameters of the nonlinearity are calculated. Predicted values of summation and difference tone levels are calculated on the basis of a power-series nonlinearity for subsequent comparison with experimentally determined subjective tone levels.

In the second experiment, masked thresholds at selected combination tone frequencies are determined with the same paradigm used in Experiment One. The existence of summation tones is demonstrated and their amplitudes estimated from the tone-on-tone masking data. The empirically obtained estimates and predicted values from Experiment One are compared. In addition, the masking of a difference tone below the primary frequencies is compared with that of the combination tones above the primary frequencies.

In the third experiment, estimates of the difference tone $2f_l - f_h$ are obtained in three ways: (1) by replication

of the cancellation paradigm (Goldstein, 1967); (2) with the cancellation paradigm corrected for possible biases of masking; and (3) by a non-simultaneous binaural loudness balance technique. The characteristics of the difference tone $2f_1-f_h$ are compared with those of the other distortion products.

Although subjective tones have some common characteristics, discrepencies between subjects will be shown for the characteristics of certain combination tones. These will be discussed in terms of the generation of combination tones and the manner in which the auditory system processes pitch.

CHAPTER TWO

AURAL HARMONIC ESTIMATION

The series of experiments presented in this chapter are partly a replication of second and third aural harmonic amplitude estimates reported elsewhere (Clack, et. al., 1972; Clack, 1975). In addition, the tone-on-tone masking data at the harmonics are examined in three other contexts: (1) is the accelerated harmonic growth rate the result of generation of the distortion product by an exponentiated term of the power series or is it an artifact of the growth of masking? (2) Are the coefficients of the terms of the power series independent of the level of the sound stimulus? If the coefficients vary with primary level then the amplitude of the distortion product will not be a simple 2dB/dB or 3dB/dB function of input level; (3) what levels of sum and difference tones would be expected if these distortion products are generated by the same nonlinearity as the harmonics?

I. METHODS

A. Subjects

Six female college students between the ages of nineteen and twenty-three served as paid observers in each of the three experiments. Prior to participating in the study they were screened for normal auditory thresholds (within 10 dB of ISO 1964 Standards) with a pulsed-tone Bekesy audiogram for each ear; they

reported a negative audiologic history. Each subject participated in one experimental session per day.

B. Stimulus Production

The method of generating the stimuli presented in the experiments is shown in Figure 2.1. All signals to the earphones were generated by low distortion oscillators (General Radio 1309) synchronized with a common reference signal originating at the system master oscillator. Tone levels were controlled by the experimenter with variable decade attenuators. In addition, the listener could control the level of the ipsilateral tone at f_{DP}, when appropriate to the experiment, with a Grason-Stadler recording attenuator. A contralateral tone at f_{DP} could be controlled with a 50 dB attenuator mounted within an enclosure which prevented the subject from acquiring visual cues relating to attenuator setting. Attenuation settings in the contralateral channel were monitored with a voltmeter (Hewlett-Packard 400 FL) and manually recorded. Stimulus sequences were controlled by two Grason-Stadler Model 471 timers which controlled Grason-Stadler Model 829 electronic switches.

All distortion products physically present in the acoustic signal were maintained at least 65 dB below the level of the generating primaries as measured with a ½" condenser microphone, 6 cc coupler, and General Radio 1900 A wave analyzer (see Table 2.1). This was accomplished in four ways:

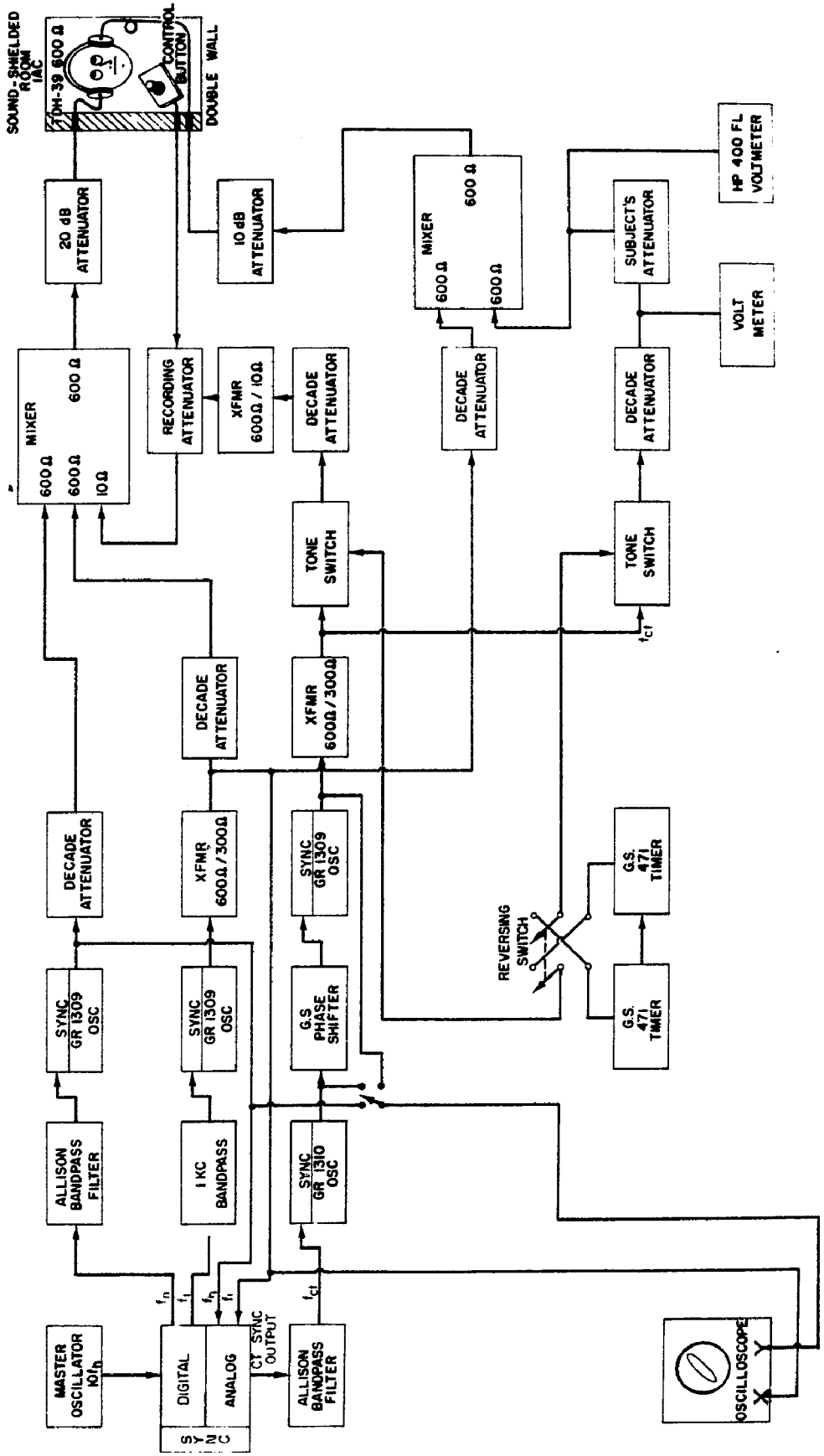


Figure 2.1: Equipment configuration for production of acoustic stimuli.

STIMULUS DISTORTION LEVELS*
dB re f_1
LEFT (IPSILATERAL) EAR

PRIMARY RATIO f_h/f_1	DISTORTION PRODUCT	f_{ct} (kHz)	SIGNAL DISTORTION dB re f_1
—	$2f_1$	2	- 65
—	$3f_1$	3	- 70
1.2	$2f_1 - f_h$	0.8	- 70
1.2	$f_1 + f_h$	2.2	- 69
1.2	$2f_1 + f_h$	3.2	- 68
1.6	$2f_h - f_1$	2.2	- 80
1.6	$f_1 + 2f_h$	4.2	- 66
1.6	$2f_1 + f_h$	3.6	- 70

*NOTE:

1) f_1 = 1 kHz
2) Measurements made at 90 dB SPL, f_1 and f_h
3) For $f_h/f_1 = 1.2$ $f_1 + f_h$ is equal in frequency to $f_h/f_1 = 1.6$ $2f_h - f_1$ = 2200 Hz

Table 2.1: Nonlinear distortion levels in the acoustic stimulus.

(1) Low distortion synchronized oscillators were used for generation of all acoustic signals.

(2) Passive linear devices were used where possible.

(3) Major sources of distortion (phase shifter and synchronizing signals) were confined to the synchronizing inputs of the oscillators. The remaining source of distortion, the electronic switch, preceded the attenuators so that distortion products were attenuated by the same factor as the signal.

(4) Distortion sources in the synchronizing pathways were bandpass filtered.

Phase synchronization of the primary tones with the tone at the distortion product frequency, was accomplished by an electronic synchronizer designed for the purpose. This unit subserves two related functions: it generates signals which synchronize oscillators producing two tones of specific frequency ratio f_{high}/f_{low} and it generates selected intermodulation products to which f_{DP} is phase-locked.

The generation of signals at f_h and f_l is accomplished digitally with two counters driven from a common source at frequency $10f_h$. One is a fixed, base 10 counter which produces a square wave at f_h. The other is a selectable base counter with range of base 12 through base 20 and

base 30. This counter produces a square wave at f_1. These signals are filtered and then used to synchronize oscillators at their respective frequencies.

The intermodulation product synchronizing generator consists of two multipliers; one in the multiply configuration and the other configured to produce the square of the input. Generation of $f_h \pm f_1$ is accomplished by bypassing the squaring operation and feeding f_h and f_1 into the multiplier. The products $2f_h \pm f_1$ and $2f_1 \pm f_h$ are generated by feeding the appropriate f_h or f_1 to the squaring multiplier than feeding the squared output and the appropriate f_h or f_1 to the following multiplier. A bandpass filter must be incorporated to reject the unwanted distortion product.

C. Aural Harmonic Estimation

Second and third aural harmonics of 1000 Hz (f_1) were estimated using the tone-on-tone paradigm detailed previously (Clack, 1967, 1968). During this aspect of the study, subjects were instructed to trace their thresholds for a pulsing acoustic signal at the distortion product frequency ($2f_1$ or $3f_1$) as in Bekesy audiometry. Each measurement sequence began and ended with the observer tracing her threshold for the appropriate pulsed (500 msec. duration, 10 msec. rise/fall) pure tone in the absence of masking. Subsequent to the first quiet tracing, a continuous tone of 1000 Hz was presented at some fixed level.

The subject was instructed to disregard this masker and continue threshold tracing. After producing a consistent response (determined by the experimenter for ten attenuation reversals) the phase of the acoustic maskee was incremented 60 degrees with respect to itself, and the tracing procedure repeated. Seven phase conditions equally spaced every 60 degrees between 0 degrees and 360 degrees were presented. Primary levels and distortion product frequency were changed unsystematically between sessions. Each condition was repeated at least once.

Sound pressure levels at which attenuation reversals occurred were converted to dynes/cm^2. The best fit of the equation

$$A_r^2 = A_e^2 + A_i^2 = 2A_eA_i \cos \phi \qquad (2\text{-}1)$$

where

A_e = amplitude of the acoustic distortion product

A_i = amplitude of the internal aural distortion product

A_r = constant phase independent threshold level

ϕ = phase angle between A_e and A_i

to the data was calculated by representing the data as a circle (polar coordinate representation of equation 2-1) and minimizing the area of a ring determined by the deviation of the data about the predictive equation (2-1). Full details of the algorithm have been published (Erdreich and Clack, 1972). In addition to calculating the level of the aural distortion product, A_i, and the masked threshold, A_r, a statistic was developed on which to base a decision concern-

ing the acceptability of the data. This acceptability criterion ratio (ACR) is described in an appendix.

D. Combination Tone Level Prediction

Assuming the aural harmonics arise from a power-series nonlinearity of the form

$$X = a_0+a_1p+a_2p^2 + \ldots+a_np^n \qquad (2\text{-}2)$$

the coefficients a_2 and a_3 may be calculated from the aural harmonic estimates. Substitution of $p = p_1 \sin \omega t$ into equation 2-2 yields expression for the level of the second aural harmonic

$$L_{AH2} = a_2p_1^2/2 \qquad (2\text{-}3)$$

and for the third harmonic

$$L_{AH3} = a_3p_1^3/4 . \qquad (2\text{-}4)$$

From these equations a_2 and a_3 are calculated by simple rearrangement of the terms and substitutions of the sound pressure and aural harmonic level in dynes/cm^2 into the appropriate equation. To predict combination tone levels arising from a power-series nonlinearity, the stimulus

$P = p_1 \sin \omega_1 t+p_h \sin (\omega_h t+\phi)$ is substituted in equation 2-2.

This yields upon simplification

$$L_{2f_1\pm f_h} = 3a_3p_1^2p_h/4 \quad \text{and} \qquad (2\text{-}5)$$

$$L_{2f_h\pm f_1} = 3a_3p_h^2p_1/4. \qquad (2\text{-}6)$$

For the quadratic term the corresponding levels are

$$L_{f_h \pm f_1} = a_2 p_1 p_h. \qquad (2\text{-}7)$$

Appropriate combination tone levels are calculated by substitution of a_2 or a_3 and the masker tone levels into equation 2-5, 2-6, or 2-7.

II. MEASUREMENTS OF A HARMONIC MASKEE IN THE PRESENCE OF A FUNDAMENTAL

A. Thresholds

Several characteristics of the measured threshold shifts have been reported (Clack, 1967, 1968; Clack, et. al., 1972). The level of the average sinusoidal variation generally increased with increasing masker level and absolute phase of the maximum and minimum maskee level varied between subjects and sessions. The maximum and minimum were separated by 180 degrees.

Figure 2-2 illustrates the masked threshold for a 2000-Hz or a 3000-Hz tone in the presence of a 1000-Hz masker for two subjects: ED, who is typical of five of the six listeners, and DY, who represents one subject with low distortion product levels. The solid curves are the fit of the predictive equation (2-1) to the threshold measures. As masker level increases, the amplitude of the sinusoidal threshold effect first increases and then with continued masker level increase, decreases slightly. The data is well fit by equation 2-1 as evidenced by the low acceptance criterion ratios.

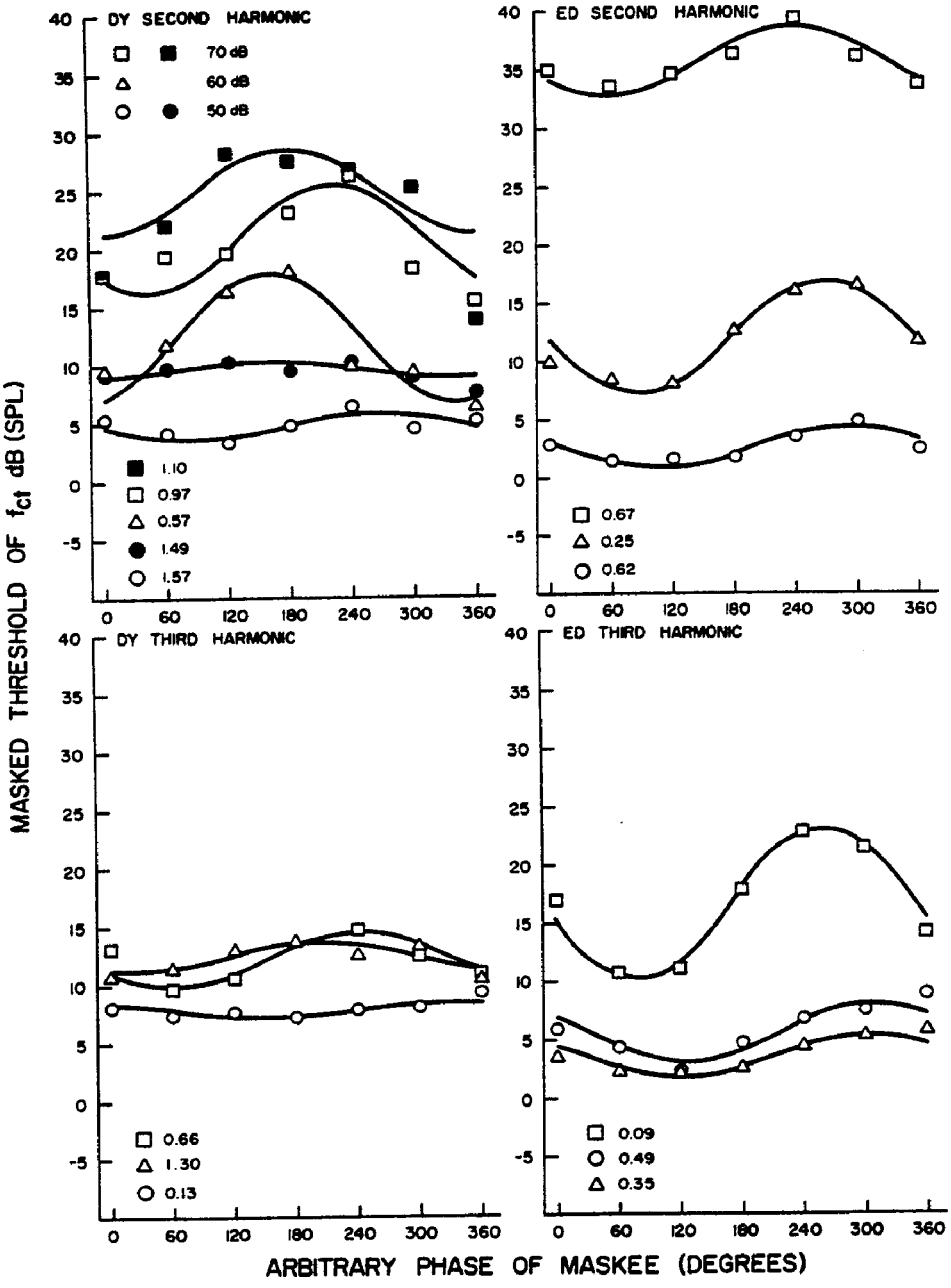


Figure 2.2: Sinusoidal threshold shifts of a 2000 Hz or 3000 Hz acoustic maskee in the presence of a 1000 Hz (M) for two subjects. The numbers in each figure are the value of the acceptance criterion ratio (ACR).

Only DY at a masker level of 50 dB SPL is excessively variable.

Calculated masking of the harmonic by the fundamental at different primary intensities increased at the same rate for all observers. Therefore, masked thresholds are presented as averages for all six listeners. These thresholds at 2000 Hz and at 3000 Hz in the presence of a 1000-Hz masker are illustrated in Figure 2-3.

For masker levels between 60 and 70 dB SPL at 1000 Hz, masking at 2000 Hz, the second harmonic, increases at approximately 2dB/dB. Below the 60 dB SPL masker level, the rate of change of masking decreases. Masking at the third harmonic of 1000 Hz increases more slowly between equivalent masker levels than masking at the second harmonic, and for equivalent masker intensities is lower at the third harmonic than at the second.

B. Distortion Products

Estimates of second and third aural harmonics for all subjects are presented in Figure 2-4. At both harmonics, DY manifests distortion product amplitudes which increase at less than 1dB/dB. From the acceptance criterion ratio for the analytic fit of the data, we inferred that these estimates were valid. Of the five remaining observers, four exhibited second harmonic growth approximating 2dB/dB. The fifth, EL, produced threshold data which at the 50 dB SPL and 70 dB SPL

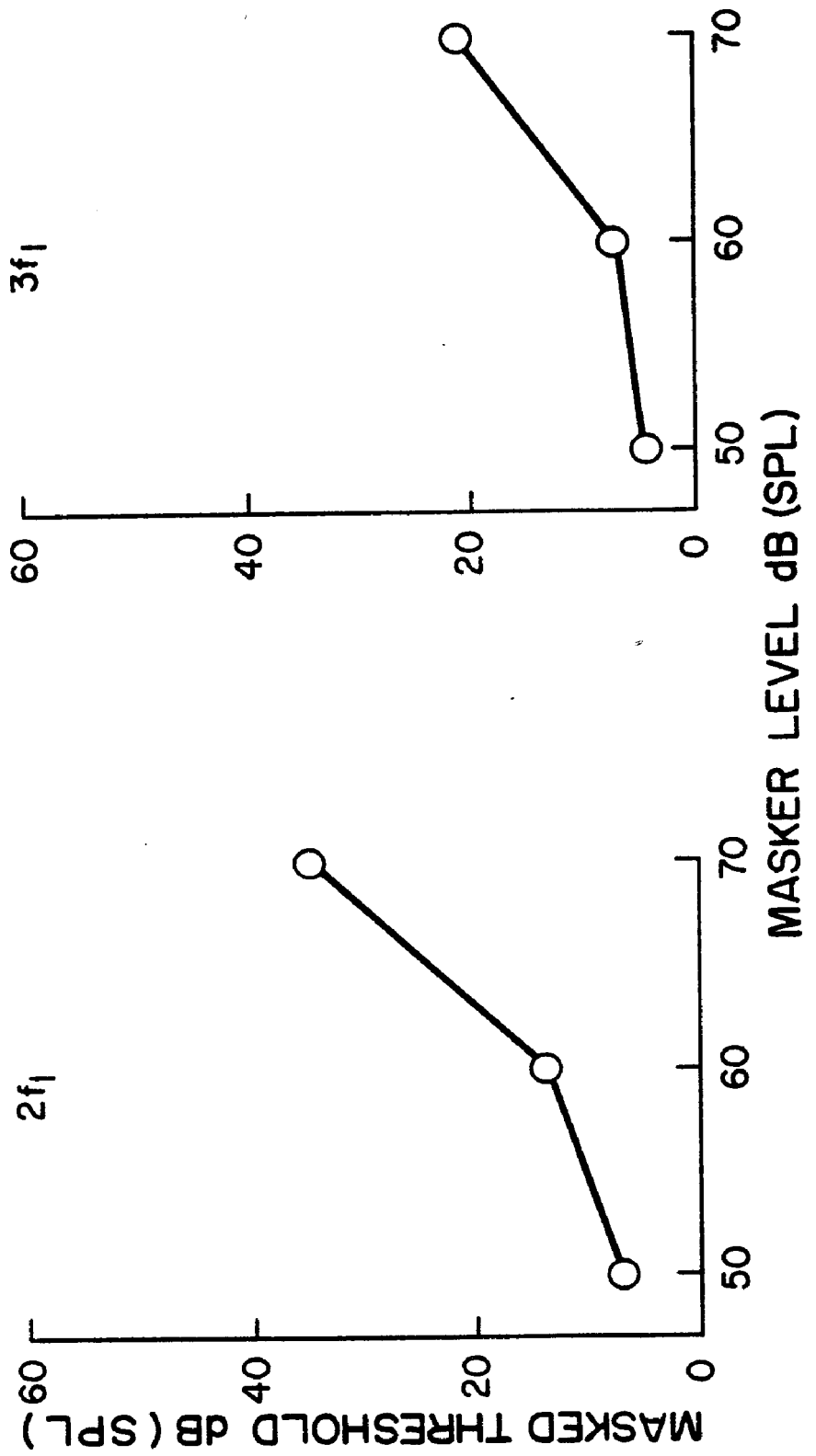


Figure 2.3: Average masked thresholds at the second harmonic (2000 Hz) and at the third harmonic (3000 Hz) of a 1000 Hz primary tone for all listeners.

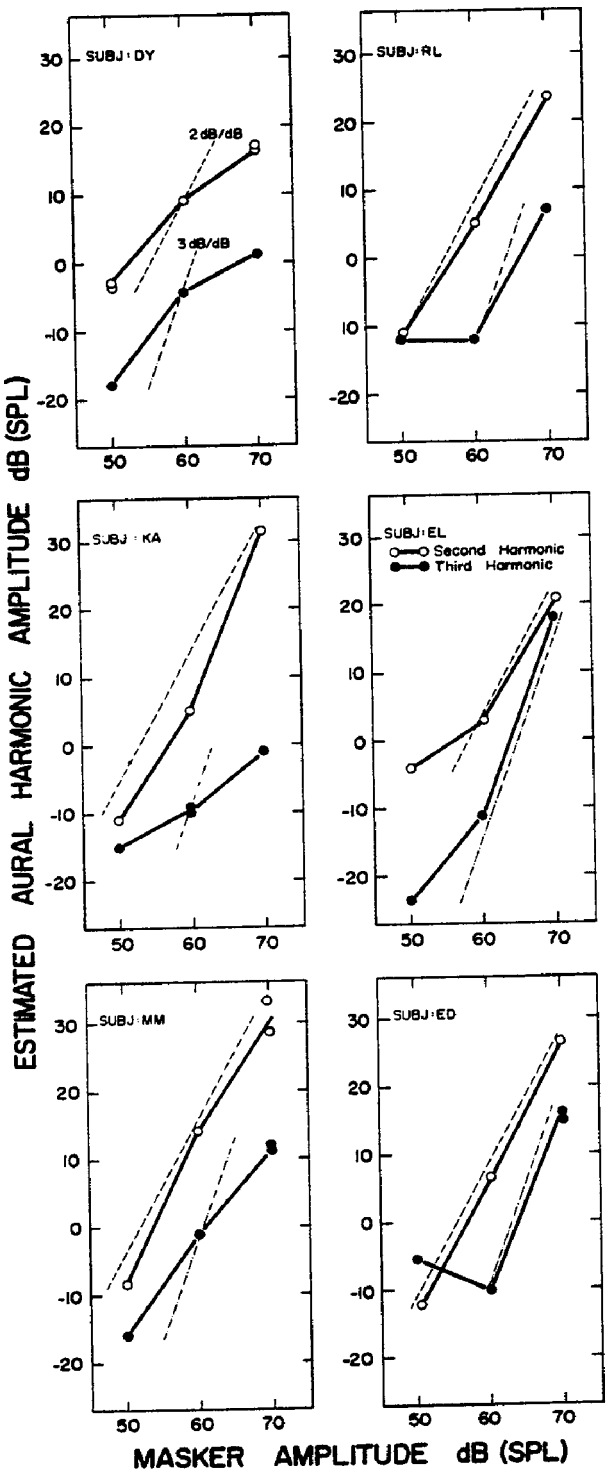


Figure 2.4: Second and third aural harmonic estimates for a 1000 Hz fundamental frequency masker.

masker conditions produced estimates of questionable validity (ACR at 50 dB = 2.3, ACR at 70 dB = 2.0).

Third harmonic growth functions demonstrated widely variable slopes ranging from approximately 3dB/dB to less than 1dB/dB. There was no consistent relationship between the slope of the growth function and estimate validity, although the three high values of the ACR were found at only lower masker levels (MM 50 dB ACR = 1.9, EL 50 dB ACR = 2.6, EL 60 dB ACR = 1.9).

Variability tends to be greatest at low masker levels, and for these masker levels the distortion product estimates determined from valid data are very low, suggesting that at low primary levels noise is of a sufficient level to limit the accuracy of the estimation paradigm.

C. Predicted Combination Tone Levels

Coefficients a_2 and a_3 of the power-series were calculated from aural harmonic estimates using equations 2-3 and 2-4. These coefficients and the cubic and quadratic distortion product levels predicted from them are presented in Table 2.2. The second order combination tones are calculated to be 6dB higher than their respective second aural harmonic; third order combination tones are predicted as 9.5 dB and higher than the corresponding third aural harmonic. These differences are a direct result of differences of ½ (second order) and 3/4 (third order) in the power series predictions for harmonic vs. combination tone levels.

AURAL HARMONIC ESTIMATES AND POWER SERIES COEFFICIENTS AT 60 dB (SPL) AND 70 dB (SPL)

SUBJECT	MASKER dB (SPL)	$\hat{AH}_2$ dB (SPL)	$\hat{AH}_3$ dB (SPL)	a_2	a_3	PREDICTED ADP LEVEL QUADRATIC dB (SPL)	PREDICTED ADP LEVEL CUBIC dB (SPL)
DY	60		-4.7		.0580		4.8
	70		0.7		.0034		10.3
	60	8.9		.0280		15.2	
	70	17.5		.0075		23.5	
	70	15.6		.0060		21.7	
EL	60		-10.7		.0291		-1.2
	70		18.3		.0259		27.8
	60	3.1		.0142		9.1	
	70	20.1		.0101		26.1	
KA	60		-9.1		.0352		0.5
	60		-10.8		.0289		-1.2
	70		-0.6		.0029		8.9
	60	4.4		.0167		10.5	
	70	26.2		.0205		32.3	
ED	60		-10.5		.0297		-1.0
	70		14.6		.0169		24.1
	60	7.1		.0227		13.2	
	70	26.4		.0210		32.5	
MM	60	14.1		.0505		20.1	
	70	28.6		.0269		34.6	
	70	33.2		.0457		39.2	
	60		-1.2		.0875		8.4
	70		11.6		.0120		21.1
	70		11.0		.0112		20.6
RL	60		-11.7		.0026		-2.1
	70		6.8		.0069		16.4
	60	4.2		.0161		10.2	
	70	23.8		.0156		29.9	

Table 2.2

Within subjects, at 60 and 70 dB SPL the coefficients a_2, and to a lesser degree a_3, are approximately equal.

III. DISCUSSION

A. The Variation of Masked Threshold with Phase

Fundamental to the distortion estimation paradigm of this experiment and that to be presented in the next chapter is the assumption that the sinusoidal threshold shifts are the result of interaction between aural distortion product and the acoustic maskee. Although there is data which supports the validity of this assumption (Clack, et. al., 1972) the monaural phase effects observed in this experiment do not contribute to the resolution of the origin of the phenomenon. Data obtained in the next experiment are directly related to this question and for this reason, discussion of the origin of the monaural phase effect will be continued after reporting of Experiment Two.

B. Growth of the Aural Harmonic

Aural harmonic amplitudes calculated from the tone-on-tone masking data are 10-20dB below threshold, explaining their inaudibility.

With the exception of the results from DY, second harmonic amplitude estimates change with masker level at approximately 2dB/dB, agreeing with results expected from the power-series nonlinearity mechanism. In contrast with the second harmonic, estimates at the third harmonic of 1000 Hz were more variable in their

rate of growth. At masker levels between 50-60 dB SPL the low growth rates of the subjective tone may be related to low signal/noise ratio with concomitant high variability. At masker levels between 60 and 70 dB SPL the third harmonic estimates from three listeners increased at 1dB/dB or less, three other listeners exhibited growth approximating 2-3dB/dB. Such variability precludes supporting either the power-series or essential nonlinearity hypotheses on the basis of the third harmonic data alone; but demonstration of aural harmonics necessitates modification of the post-nonlinearity filter proposed by Goldstein (1967). The demonstrated existence of the harmonics precludes the existence of a filter which eliminates the high frequency subjective tones of this class. The inaudibility of the harmonic is explained on the basis of masking by the primary.

C. Masking at the Harmonic

With the masked threshold data at the second and third harmonic reported in this study we can show that the aural harmonic growth rate is not an artifact of masking. If this were true, both the second and third harmonic would increase at the same rate as the masking at the harmonic. Although the second aural harmonic and masking at 2000 Hz do follow a parallel course, this is certainly not the case with the third harmonic. In fact, this distortion product grows at a rate usually faster than the second harmonic but mask-

ing at 3000 Hz increases more slowly than masking at 2000 Hz.

D. Masker Level Dependence of Parameters of the Nonlinearity

The changes in distortion product level with primary amplitude changes may result either from changes in the coefficient of the term of the nonlinearity responsible for a particular distortion product (a_2 or a_3), from the change in the input to the nonlinearity, or a combination of both.

The agreement within listeners of calculated values of a_2 or a_3 at two intensities of the primaries suggests that the nonlinearity is intensity independent; that is, the characteristics do not change with intensity and the distortion levels are a simple function only of stimulus amplitude.

CHAPTER THREE

ESTIMATION OF COMBINATION TONES WITH THE TONE-ON-TONE MASKING PARADIGM

This chapter deals with the crucial missing evidence needed to compare nonlinear mechanisms; the existence of the summation tones. Both the theories of Helmholtz and of Fletcher predict harmonics, difference tones and summation tones; whereas Goldstein's theory acknowledges only difference tones below the frequency of the primaries. Furthermore, if the monaural phase effect exists at summation tone frequencies, we must ask which of the postulated mechanisms of its generation are valid for summation tones. Do any short-term phase relationships exist between the primaries and maskee which can explain the phase effect? Are there any envelope variations on which detection may be based which would result in the sinusoidal threshold shift? Last, is vector summation between the maskee and the distortion product a plausible explanation of the phase effect?

The same question of the rate of growth of the combination tones and its relation, if any, to masking which was asked of the harmonics needs to be answered for the combination tones - is the growth real or is it an artifact of masking? Another point related to the previous experiment concerns the mechanism of production of the summation tones.

If the same power-series nonlinearity which produces the

harmonics generates the combination tones, then the predicted combination tone levels calculated in the first experiment should agree with those measured in this experiment.

If the combination tone and harmonics are produced by the same power-series nonlinearity can a distinction be made between the Helmholtz model and the Fletcher model based on the combination tone estimates? Specifically, if Helmholtz' model of the nonlinearity is valid, there should be differences in the amplitudes of particular combination tones when generated by different frequency primaries; these differences would not exist if the Fletcher model is valid. Furthermore, with Helmholtz' model relative growth rates of combination tones predicted from the power-series calculated in the former experiment are expected to be accurate,but the absolute values of the distortion products would be in error. Conversely, if the Fletcher's model is correct, the absolute as well as the relative values of the subjective tones would be independent of the frequencies of the tones or the primaries; the predicted summation tone levels from Experiment One should agree with measured values in this chapter.

The measured and predicted distortion product levels will be compared and the effect of frequency separation between primaries shown after presentation of the tone-on-tone masking data obtained in this experiment.

I. METHODS

The paradigm described in Experiment One was used for estimation of combination tones $2f_1+f_h$, $2f_h+f_1$, and in some cases for $2f_1-f_h$ and f_h-f_1. The only modification was that two phase-synchronized masking tones at a frequency ratio of $f_h/f_1 = 1.2$

or 1.6 were used.

II. MASKED THRESHOLD MEASUREMENT ABOVE THE FREQUENCY OF THE PRIMARIES

A. Monaural Phase Effects

Reliable threshold shifts were found with masker levels as low as 40 dB SPL. Threshold data at $f_{DP} = 2f_1+f_h$ and $f_{DP} = f_h+f_1$ is shown in Figure 3.1 for the two listeners whose harmonic data was shown in the previous experiment. At the summation tone frequencies, masked threshold for DY approximate those of ED as evidenced by the absolute level of the average sinusoidal thresholds. Other characteristics of the monaural phase effect are consistent with those reported in Experiment One.

Increasing the frequency separation between the two primaries ($f_h/f_1 = 1.6$ compared to $f_h/f_1 = 1.2$) does not change the form of the phase effect (Figure 3.2). Differences between the phase effect amplitude for the combination tone $2f_1+f_h$ appear to be limited to changes in average level (compare $2f_1+f_h$ = 3200 Hz and $2f_1+f_h$ = 3600 Hz in Figure 3.2). Similar threshold behavior is observed for $2f_1+f_h$ and $2f_h+f_1$. In contrast, the average threshold values for $2f_h-f_1$ are higher than those for the summation tones in the figure.

B. Masked Thresholds

Masked threshold levels for the observers similarly increased with increasing masker levels. Consistent with the presentation of masking at the aural harmonics, growth of masking at 2200 Hz (f_h+f_1, $f_h/f_1 = 1.2$) and at 3200 Hz

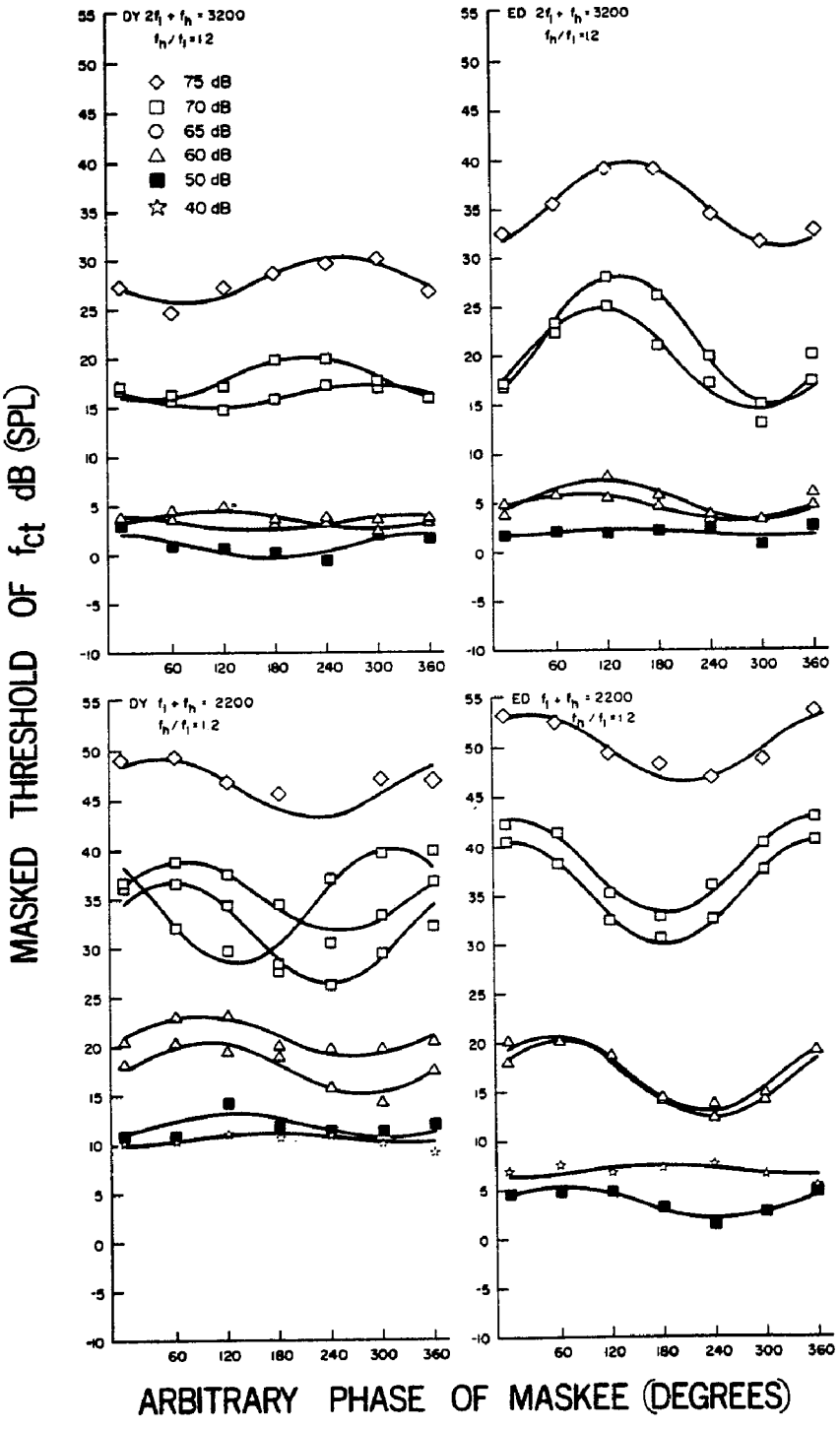


Figure 3.1: Masked threshold for cubic ($2f_1+f_h$) and quadratic (f_h+f_1) summation tones for two subjects at $f_h/f_1 = 1.2$, $f_1 = 1$KHz.

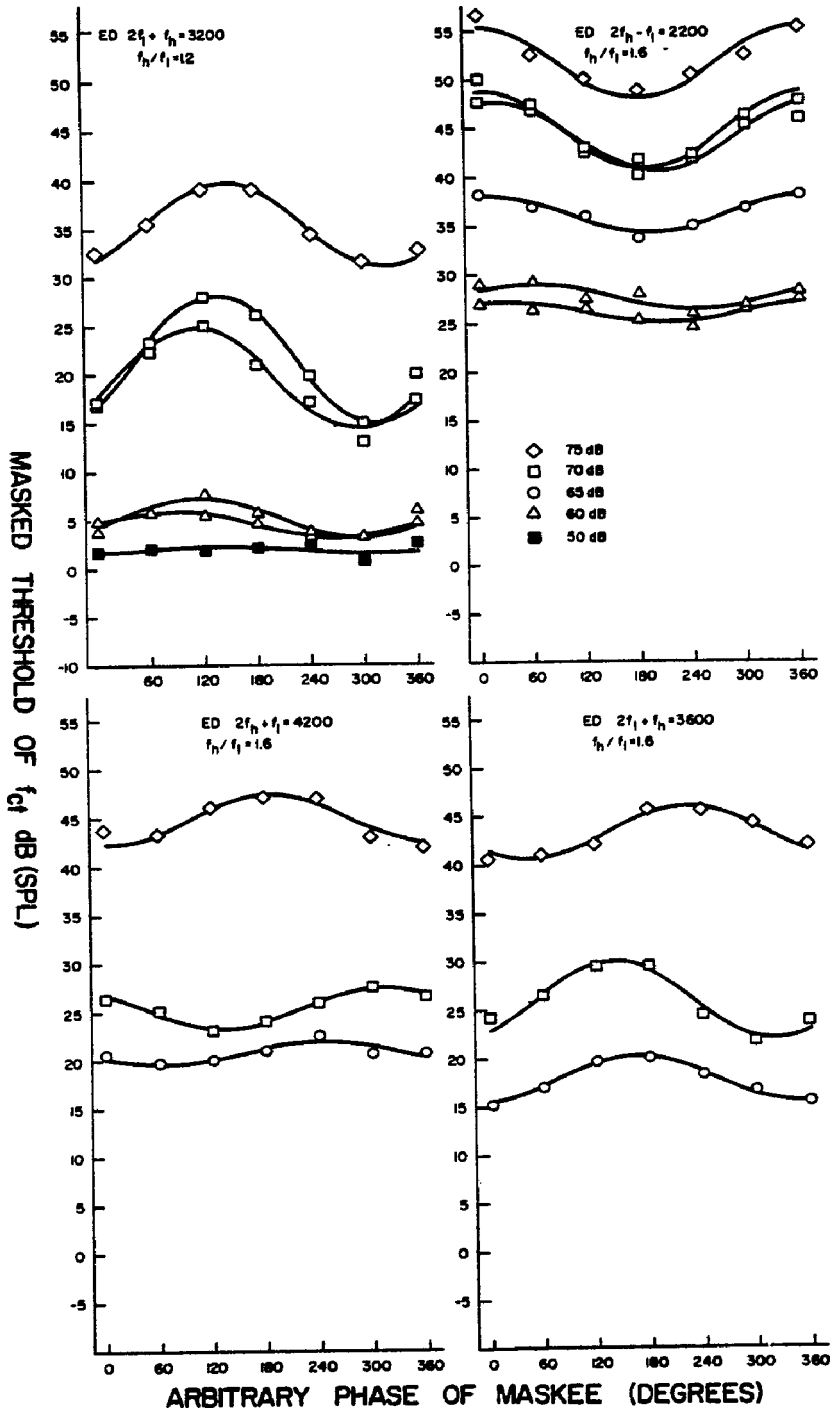


Figure 3.2: Threshold of the acoustic maskee in the presence of f_1 = 1000 Hz and f_h = 1600 Hz as phase between the acoustic maskee and aural distortion product is varied. Measures of $2f_1+f_h$ at f_h/f_1 = 1.2 are included for comparison.

($2f_1+f_h$, f_h/f_1 = 1.2) is shown in Figure 3.3 averaged for all subjects. Averaged masking data at $2f_h+f_1$ and $2f_1+f_h$ for f_h/f_1 = 1.6 is shown in Figure 3.4. At 2200 Hz, the quadratic sum tone, (Figure 3.3) there is no change in masking with changes in primary level below about 50 dB. Above this level masking grows at 2dB/dB. The same is true of the masking at 3200 Hz, the cubic sum tone, above 60 dB SPL. Masking by the two primaries at 3200 Hz was approximately 15 dB less than masking at 2200 Hz.

C. Summation Tone Estimates

The amplitude of the first-order summation tone (f_h+f_1, Figure 3.5) generally increases at an accelerated rate approximately 2dB/dB with increases of primary amplitude. At primary levels below 50 dB SPL (60 dB SPL for DY) the rate of change of distortion product amplitude shows considerable variability, in one case (KA) decreasing with an increase of primary level. In most cases, this departure from the 2dB/dB growth function is associated with a high value of the acceptance criterion ratio and therefore, may be an artifact. This is also true for the departure from 2dB/dB between 70 and 75 dB SPL for MM (ACR = 3.5).

Predicted values of f_h+f_1 determined in Experiment One and plotted in the figure agree closely with the experimentally determined estimates. Major departures are seen for predicted values determined from harmonic estimates at lower masker levels at which the problem of estimate reliability becomes significant.

Estimates of the second-order sum tone of $2f_1+f_h$ (Figure

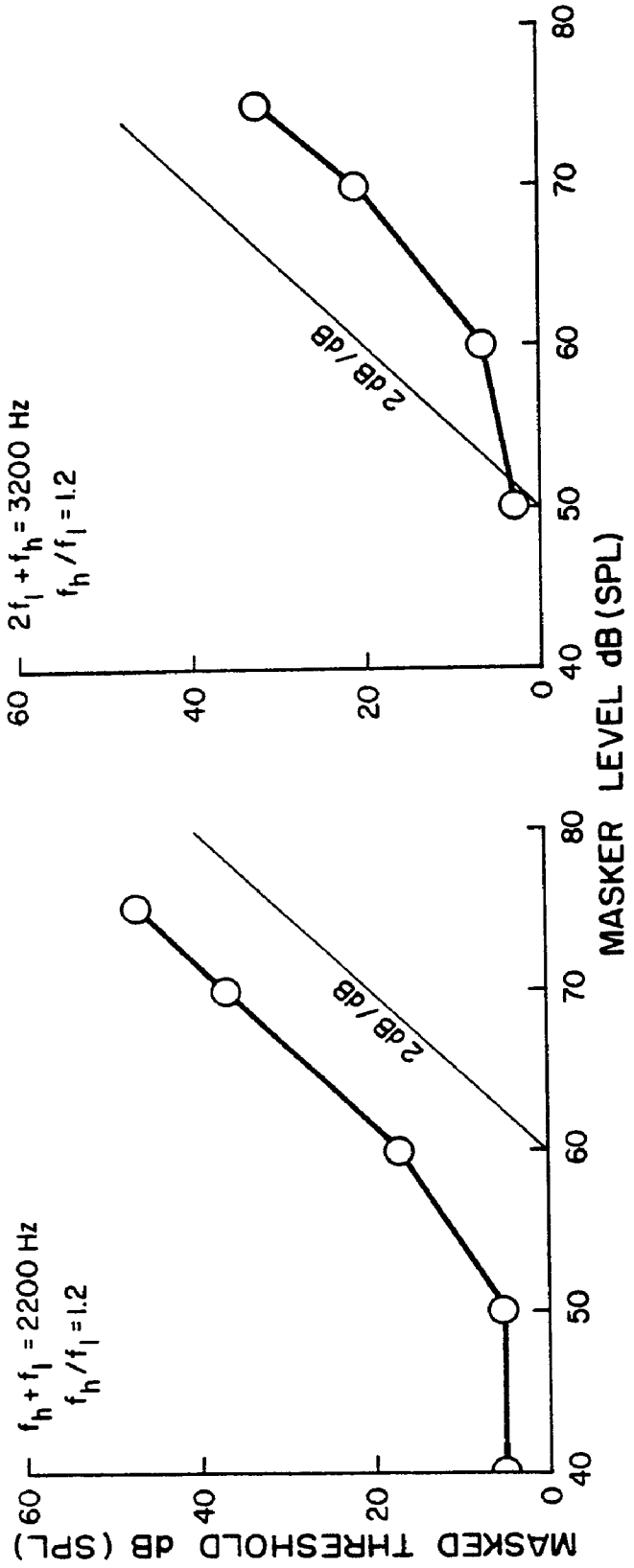


Figure 3.3: Average masked thresholds at 2200-Hz and 3200-Hz in the presence of 1000-Hz plus 1200-Hz maskers.

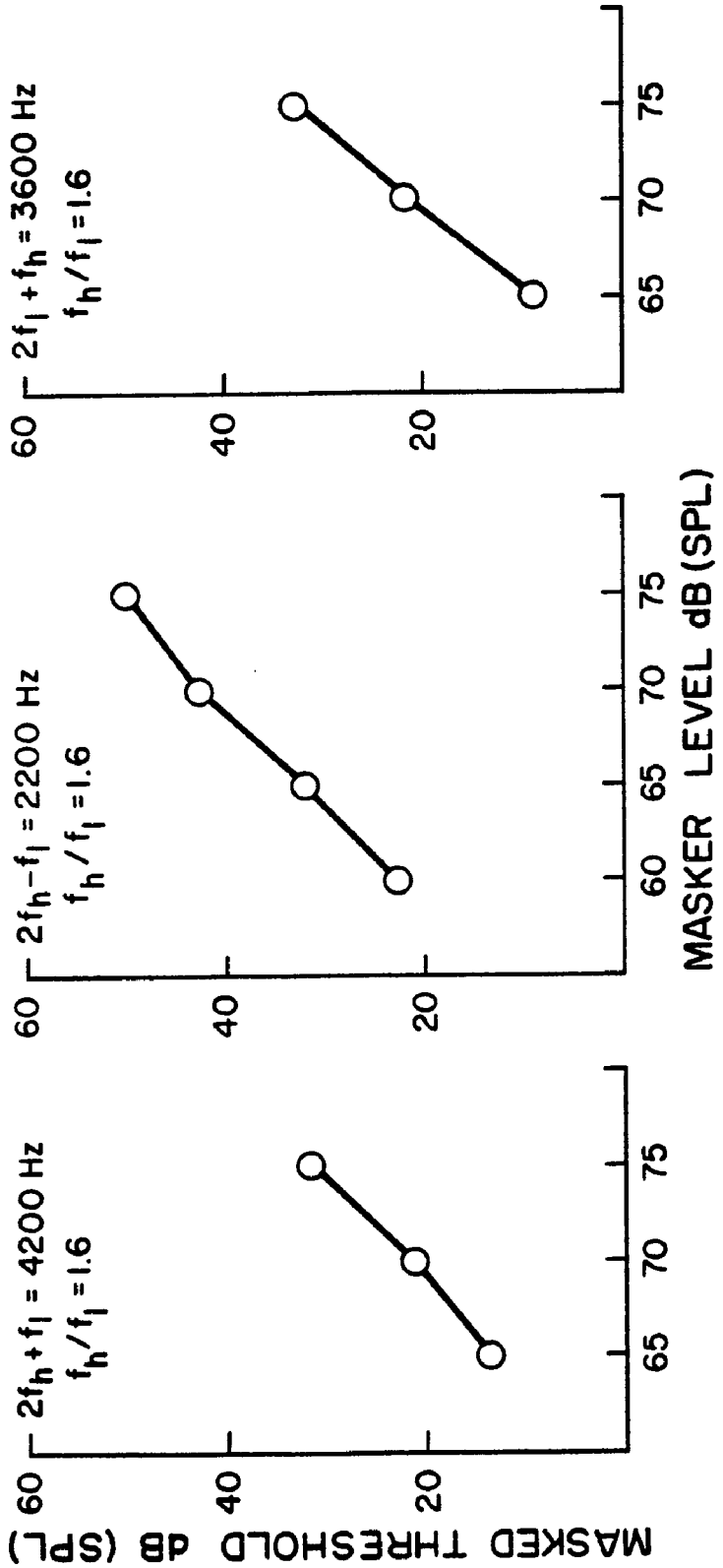


Figure 3.4: Average masked thresholds at 4000-Hz, 2200-Hz and 3600-Hz in the presence of 1000-Hz plus 1600-Hz maskers.

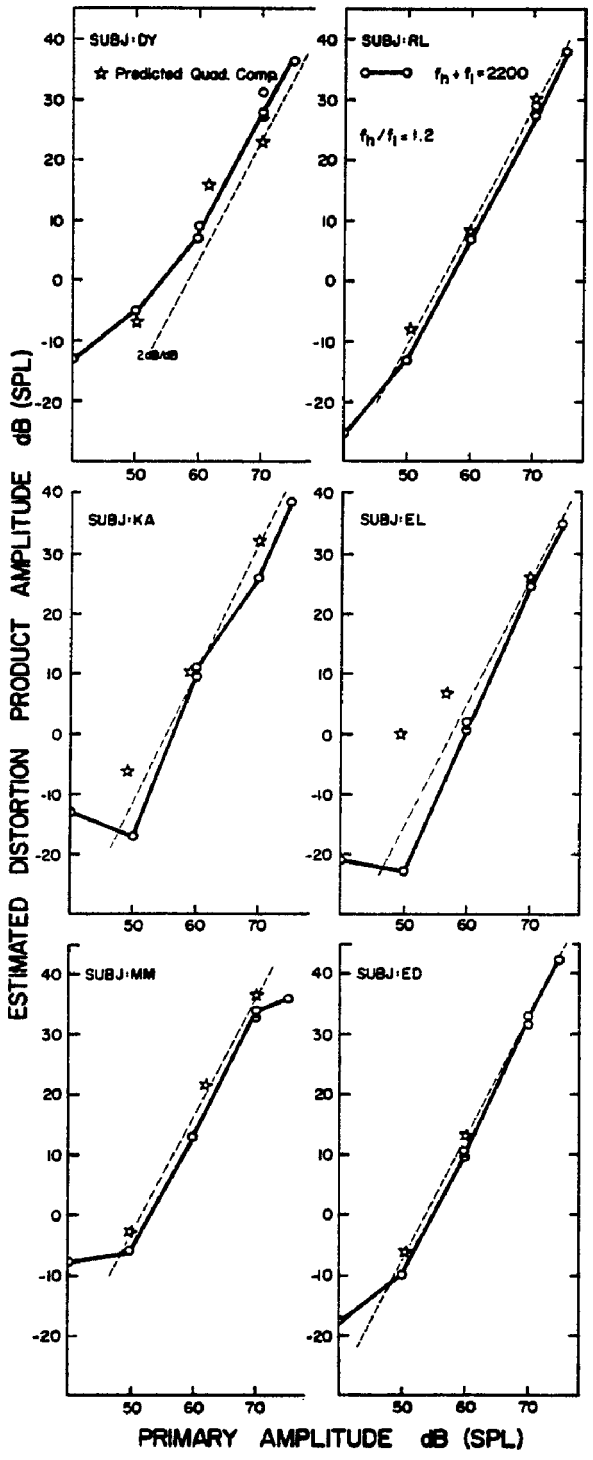


Figure 3.5: First-order (f_h+f_l) aural distortion product amplitudes as a function of increasing masker levels. Predicted values are shown by "+". Primary ratio f_h/f_l = 1.2.

3.6) manifest properties similar to those for f_h+f_1. Above some primary level the distortion product changes at an accelerated rate, in this case approximately 3dB/dB. At low levels, the rate of change decreases and the threshold data from which the estimates are made exhibit high variability. Only one exception to this trend exists for the data at $2f_1+f_h$ (KA at 75 dB SPL, ACR = 0.7).

The lack of close agreement between the predicted and observed subjective tone estimates at the lower primary levels are probably the result of limitations of third harmonic determination with the tone-on-tone masking paradigm. At 70 dB SPL, and for some listeners at 60 dB SPL, the predicted and observed values are in close agreement. Tone-on-tone masking estimates of the cubic intermodulation product appear to be reliable at masker levels lower than those for the third harmonic; combination tone estimation may be a more sensitive measure of subjective tone generation in the auditory system.

Estimated subjective tone amplitudes were determined for $2f_h-f_1$ and $2f_1+f_h$ with a primary frequency ratio of $f_h/f_1 = 1.6$ (Figure 3.7). With the exception of the estimates determined for DL, the growth and amplitudes of subjective tones $2f_1+f_h$ and $2f_h+f_1$ are similar within subjects. The points which deviate from this similarity are those with relatively high ACR's (MM at 75 dB SPL $2f_1+f_h$, ACR = 1.5; EL at 65 dB SPL $2f_1+f_h$, ACR = 1.6). At $2f_h-f_1$ (2200 Hz), the distortion product estimates are 15-30 dB higher than those at the summation tone frequencies.

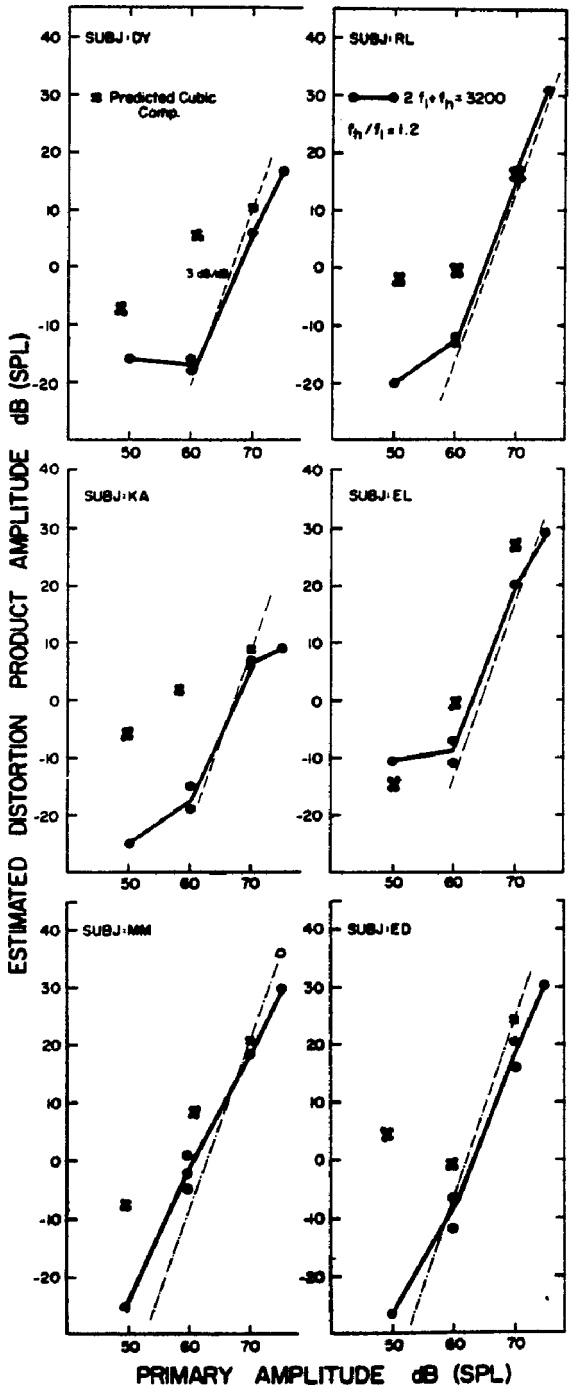


Figure 3.6: Second-order ($2f_1+f_h$) aural distortion product amplitudes as a function of increasing masker levels. Predicted values are shown by "+". Primary ratio $f_h/f_1 = 1.2$.

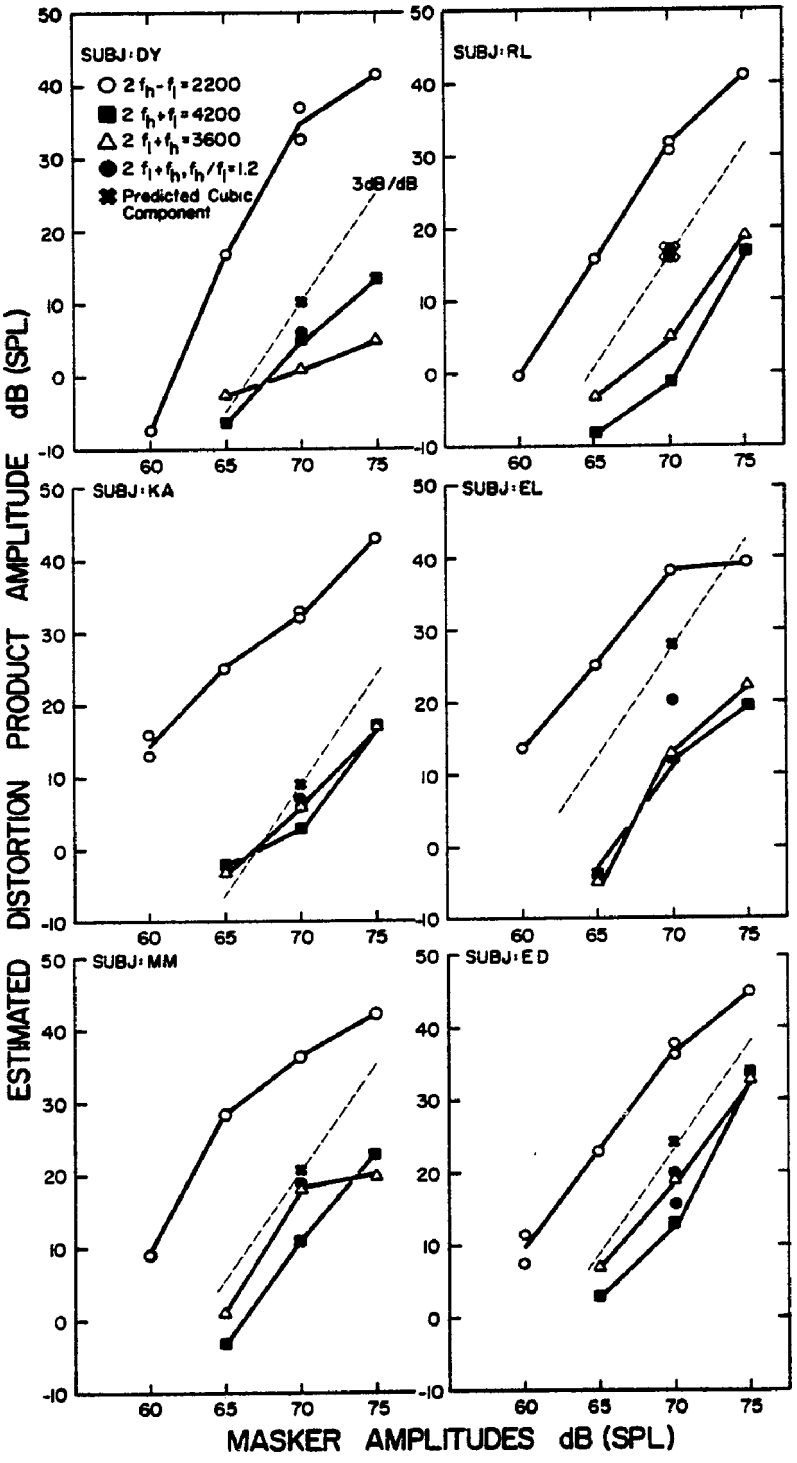


Figure 3.7: Aural intermodulation distortion product amplitudes measured at f_h/f_l = 1.6, f_l = 1KHz. Broken lines are predicted ADP levels.

Distortion product estimates increase with masker level at a rate exceeding the rate of increase of the maskers. Two points at $2f_h-f_l$ are considered invalid on the basis of the ACR: at 75 dB SPL, EL produced threshold data with an ACR = 1.9. At 60 dB SPL, threshold data for DY at $2f_h-f_l$ yielded an ACR = 3.8.

The levels of both the summation and difference tones were different than those predicted (shown by an "X" in Figure 3.7) from third aural harmonic estimates at 70 dB SPL. As previously discussed, the differences are attributable to problems with measurement of third aural harmonics.

Over the restricted range of masker amplitudes within which estimates were obtained, summation tone amplitude estimates increased with masker intensity at between 2dB/dB (KA) and 3.4dB/dB (MM, $2f_l+f_h$). The one exception to this trend was DY at $2f_l+f_h$ which increased at only 0.7dB/dB. The correspondence with the predicted 3dB/dB growth can be seen by comparison of the increase in distortion product level and the dashed lines in Figure 3.7 which slope at 3dB/dB. Similar growth characteristics are exhibited by the difference tone estimates except that the distortion product amplitudes were about 25dB higher than predicted.

D. Effects of Frequency Separation

Sensitivity of distortion product amplitude to a change in the frequency of the high frequency masker

was examined for all six listeners. The estimated distortion product amplitude for $2f_1+f_h$ (f_h = 1200 Hz or 1600 Hz) with 70 dB SPL maskers is shown in Figure 3.8. Differences between subjective tone levels with increased f_h frequency may be either positive or negative. Although the mean change is 4.0 dB (KA -10dB; EL -8.0dB; ED+2.0dB; MM 0.0dB: RL -11.5dB; DY -4.0dB) the Binomial Test (cf. Siegel, 1956) indicates that the trend is not significantly different from zero (the probability of 4 out of 5 differences being negative is 0.156). Therefore, there is no evidence which supports a frequency sensitive component of the nonlinear amplitude characteristic.

III. ESTIMATES OF f_h-f_1 WITH TONE-ON-TONE MASKING

Estimation of the quadratic difference tone f_h-f_1 with the tone-on-tone masking paradigm was attempted on subject DY. Maskee threshold variations as a function of phase are shown in Figure 3.9. The sinusoidal nature of the threshold variations and the close agreement between the data and the theoretical function (solid curve) is indicative of the distortion product being below threshold for this subject. Furthermore, the clustering of the phase amplitude effect around 35 dB (SPL) as masker amplitude (the parameter on each curve) increased indicates that the masked threshold is little influenced by a masker of higher frequency as is commonly found for pure tone masking (Wegel and Lane, 1924). The variation in masked threshold at the 200-Hz maskee

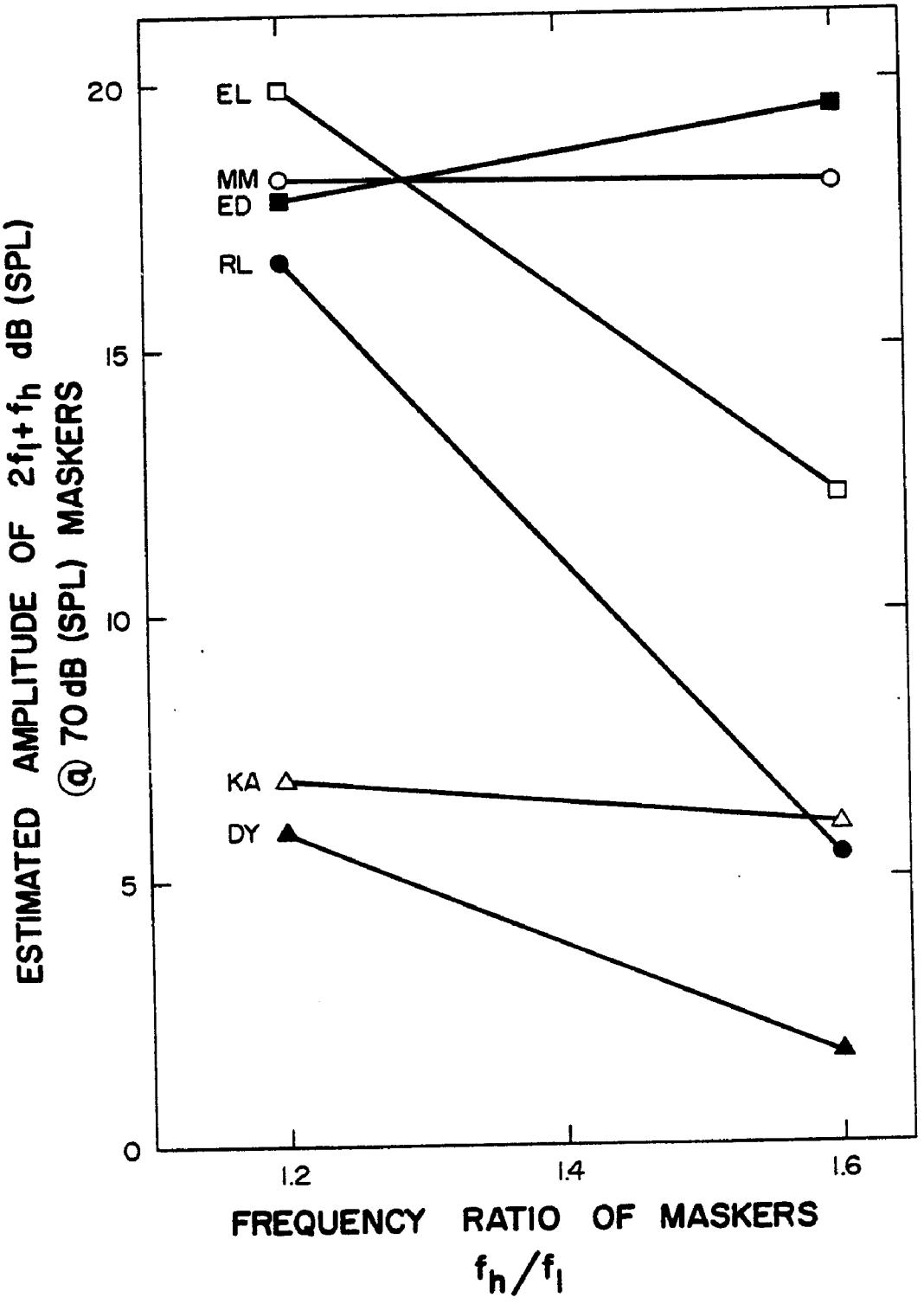


Figure 3.8: Change in amplitude of the distortion product $2f_1+f_h$ as frequency of f_h is changed from 1200 Hz to 1600 Hz. Primary amplitudes are equal to 70 dB SPL.

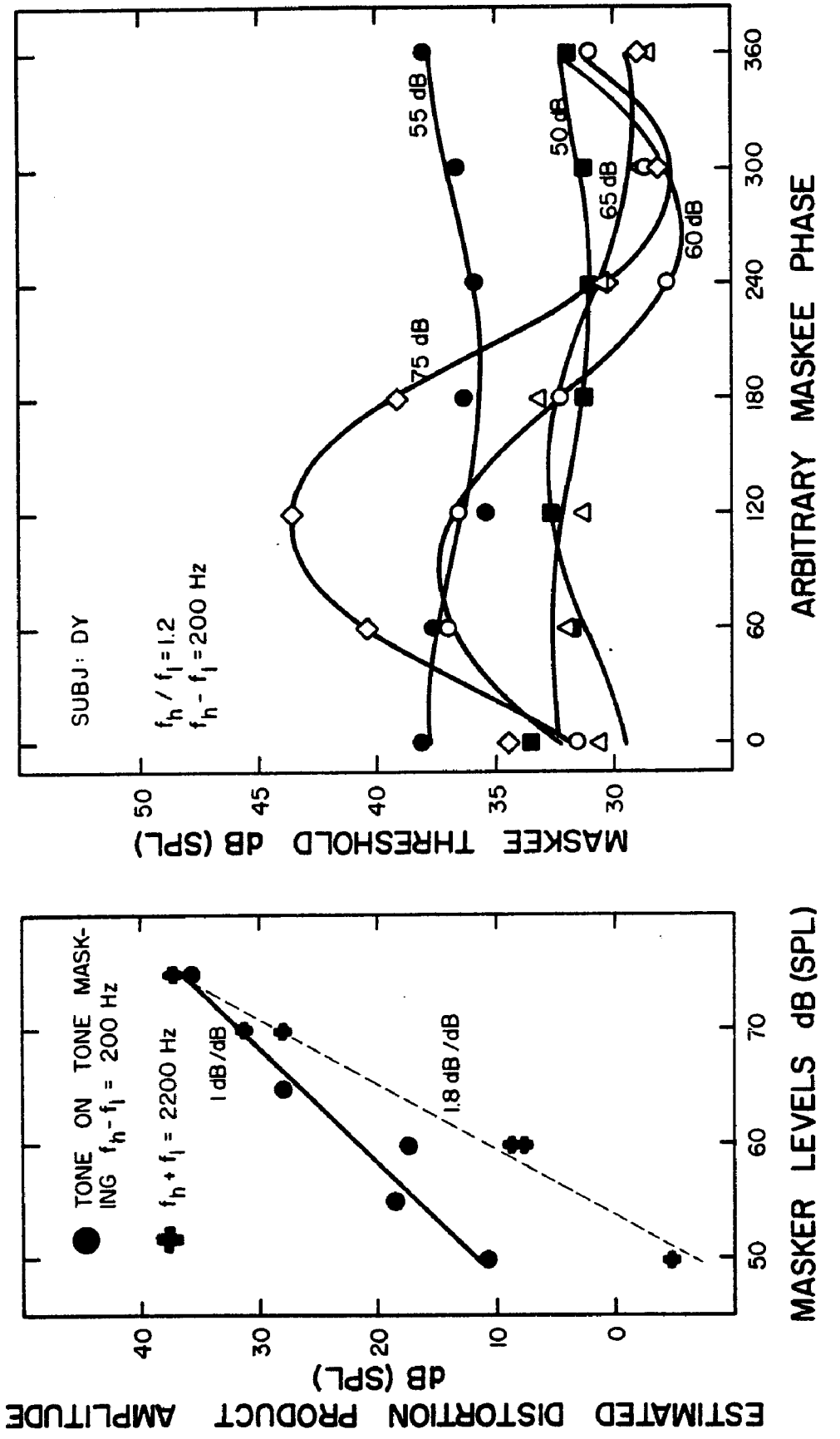


Figure 3.9: Right: Phase amplitude effects for DY tracing 200 Hz (f_h-f_l) threshold in the presence of 1000 Hz + 1200 Hz. Curve parameter is level of each of the primaries. Left: Results of tone-on-tone estimates of quadratic summation tone (QST) obtained from same subject.

frequency was between 31.0 and 38.8dB SPL over a 20 dB range of masker level.

Difference tone (f_h-f_l) amplitudes grew at only 1dB/dB in comparison with the companion summation tone which increased at 1.8dB/dB as shown on the left side of Figure 3.9.

IV. DISCUSSION

We believe that the observation of monaural phase effects at the combination tone frequencies examined in Experiment Two represents the first demonstration of summation tones since they were postulated by Helmholtz. The monaural phase effect at combination tone frequencies is evidence itself that these threshold shifts are the result of subjective/acoustic tone interaction and are not derived from phase detection of components of the stimulus.

The phase amplitude effect, to be a result of detection of relative phase between components, requires that for a given cycle of the lower frequency tone the two signals be in some static phase relation. Only in such a case can a change in phase between the signals be appreciated as a change in amplitude of one or the other signal (cf. de Boer, 1961). If the two tones, on a successive cycle basis, are not in a fixed phase relation, the relative variations between successive cycles will obscure any phase-dependent effects.

That the phase effects in tone-on-tone masking are not attributable to phase detection can be appreciated by consideration of the three signal frequencies involved in

the paradigm. In illustration, determination of $2f_1+f_h$ for $f_h/f_1 = 1.2$ involves $f_1 = 1000$ Hz, $f_h = 1200$ Hz and $f_{DP} = 3200$ Hz. Since the greatest common frequency to all three signals is 200 Hz, phase relations between the signals repeat every 5 msec. Thus, there are five full cycles of f_1, six of f_h, and sixteen of f_{DP} before the sequence repeats. Obviously, the instantaneous phase relations are not constant precluding phase variation between components as a foundation for the observed phase effects.

Other studies have also suggested that the monaural phase effects are not the result of waveform interaction or envelope detection: Chapin and Firestone (1934), Lewis and Larsen (1937), Trimmer and Firestone (1937), Clack (1967), and Clack, et. al., (1972).

Chapin and Firestone (1934) in their study of effects of phase between harmonically related tones employed a stimulus comprised of a fixed phase 104 dB SPL 108 Hz tone and a single fixed-level harmonic of 108 Hz (216, 324, 432, 540, or 648 Hz) variable in phase. Listeners were required to report any change in loudness or quality of the complex as the phase of the harmonic was varied. All of their subjects reported loudness maxima and minima corresponding to distinct phase settings separated by 180 degrees. The authors concluded that the harmonics of a fundamental are "...in phase as cosine functions without epoch angles"; perhaps more importantly, they demonstrated that subjective harmonics may constructively or destruct-

ively interfere with pure-tones at the same frequencies. At these phase settings subjects noted loudness changes in the complex as the level of the harmonic increased. Two distinct patterns were observed: for maximum loudness phase, increases in the harmonic caused increases in loudness. Level increments for minimum loudness phase resulted first in a decrease and then in an increase in the loudness of the stimulus.

Thus, the loudness appears to be proportional to the amplitude expected on the basis of summation between maskee and distortion product.

Lewis and Larsen (1937) presented data which may be interpreted as demonstrating the same vector summation for the difference tone f_h-f_1 shown by Chapin and Firestone for the harmonics. They examined the threshold change for a 130 Hz phase variable tone masked by a two-tone complex consisting of 370 Hz plus 520 Hz. Strong sinusoidal changes in threshold for the 130-Hz difference tone occurred as a function of the phase of the 130-Hz tone. In a second experiment, the phase of 130-Hz signal was set at the appropriate point for maximal destructive interference and its intensity increased. Two listeners were required to determine which of two conditions produced the louder perception of a 130-Hz fundamental: the 390/520 Hz complex alone or the complex with an added 130-Hz component in cancellation phase. At low and high levels of the added 130-Hz tone, the percentage of judgements that the stimu-

lus containing the 130-Hz component produced a louder sensation of the fundamental was higher than for intermediate levels indicating, as the authors conclude, that this is evidence for the vectorial summation mentioned above.

Clack, et. al., (1972) present data which lends further support to the vector summation of maskee and aural distortion product being responsible of the monaural phase effect rather than waveform interaction. The demonstration that phase dependent threshold shifts for a maskee tone of 1500 Hz, 2000 Hz, or 2500 Hz occur in the presence of a 1000 Hz masking tone. These threshold shifts have different character as a function of the maskee frequency. At 2000 Hz, the second harmonic of the masker, the threshold shifts are clearly sinusoidal and larger than those non-sinusoidal shifts observed at 1500 Hz and 2500 Hz. Additionally, at a masker level of 80 dB SPL narrow band noise at 500 Hz, the difference tone frequency, reduced the threshold variation for only the 1500 Hz and 2500 Hz maskee conditions. This suggests that temporal factors related to the common 2msec. (500 Hz) period of the stimulus components are not responsible for the monaural phase effect at 2000 Hz. The simplest explanation remaining is the vector summation between aural distortion product and maskee.

Based on the demonstration of the phase effects at combination tone frequencies as well as earlier studies we consider the monaural phase effect evidence of aural distortion products.

Masking of a tone at the distortion product frequency by the primaries follows patterns known for masking of one tone by another (Egan and Hake, 1951). Masked threshold for distortion product frequencies higher than the primaries increased at 2dB/dB independent of the order of the combination tone. Since the combination tones were found to grow at rates dependent on their order - rates different from the masking growth - the distortion producing mechanism must be independent of auditory masking.

In general, the combination tones higher in frequency than the primaries exhibited an accelerated growth rate. Departures from 2dB/dB or 3dB/dB slopes are usually in the direction of reduced slope. Reasons for these differences are not yet clear.

All subjects occasionally report that the quality of the stimulus complex changes with variations in the phase of f_{DP}. The listener, instructed to trace threshold for the pulsed f_{DP} is thereby given an additional parameter on which to base the decision, "I hear the tone." If at subsequent phase settings the listener delays his decision because she is waiting for the perception of the previous quality of the complex this may reduce the level at which the maskee is detected, in turn decreasing the apparent growth of the distortion product.

Additionally, contributions of higher ordered terms of the nonlinearity from different areas of cochlea (Dallos, 1973) or different order distortion products of equal

frequency (Hall, 1972a) may combine to increase or diminish the subjective tone obscuring the true form of distortion product generation.

What is most important is that the subjective tone levels grow at an accelerated rate and that these rates are clearly similar to those predicted from the power-series hypothesis. Departures from precisely predicted rates are small and apparently the result of secondary effects not directly relating to the question of the form of the nonlinear mechanism.

The agreement between predicted and observed distortion product growth rates combined with the agreement between predicted absolute levels of combination tones for primary ratios of $f_h/f_1 = 1.2$ is evidence of the identity of the mechanisms producing the aural harmonics and the summation tones. Similarly, the equality of subjective tones $2f_1+f_h$ and $2f_h+f_1$ at $f_h/f_1 = 1.6$ suggests that these are distortion products from a nonlinearity of the same form with equal third order coefficients, i.e., the same nonlinearity. The high relative level of $2f_h-f_1$ with growth rates in the range of 3dB/dB may be related to the mechanism by which the distortion product, once generated, is processed by the auditory system.

Psychophysical evidence of cancellation of a subjective tone by a pure tone of equal frequency (Goldstein, 1967; Hall, 1972; Smoorenburg, 1972), the vectorial summation of distortion product and maskee in tone-on-tone masking, and

the reduction of neural synchrony in response to an appropriately adjusted cancellation tone demonstrated by Goldstein and Kiang (1968) all suggest that the combination tone, once generated, behaves much as if it were an externally introduced pure tone.

If the degree of excitation at the distortion product "frequency place" is dependent on both the frequency (i.e., the locus) of the primaries and the locus of excitation by the combination tone frequency, then subjective tones closer to the primary frequencies will act more strongly than those further removed from the place at which the distortion product is localized. Since the cubic difference tone $2f_h-f_1$ is closer in frequency to the primaries than are the summation tones, the hypothesized frequency dependent weighting of the recoupling causes them to appear greater in amplitude than the equivalent sum tones. Goldstein (1967) and Smoorenburg (1972b) have suggested a frequency selective weighting operating on both the primary interactions and the distortion product recoupling. This may explain variations in the effectiveness of the stimuli which are responsible for differences between predicted and observed summation tone levels for $f_h/f_1 = 1.6$.

If the aural harmonic frequencies on which predicted combination tone levels are based approximate the combination tone frequencies, then decreased excitation of one distortion component will be offset by decreases in the other and the predictions will agree with observations.

Conversely, if decreases in excitation caused by the recoupling of the distortion products are not equal, the tone which is more efficiently transferred will appear of higher level than that further removed from the locus of the primaries. If it is assumed that the excitation by combination tones spread in the same manner as masking by a pure tone, and furthermore that the spread to higher frequencies decreases at approximately 20-30 dB/oct. as shown by Ehmer (1959), then the levels of $2f_h+f_1$ (4200 Hz) and $2f_h+f_1$ (2200 Hz) ($f_h/f_1 = 1.6$) which are created equal in amplitude will have effective levels approximately 25 dB different as observed for our listeners. Furthermore, according to this theory, the 3600-Hz summation tone should be slightly greater than the 4200-Hz summation tone as was observed in five of the six listeners.

CHAPTER FOUR

MEASUREMENT OF THE AUDIBLE DIFFERENCE TONE $2f_1-f_h$

Although the results of the previously reported experiments tend to support the power-series nonlinearity model, the results reported by Goldstein (1967) remain unexplained. The only model consistent with Goldstein's measurements is an essential nonlinearity. Now that the cubic and quadratic sum tones have been demonstrated we must ask if it is plausible that the difference tone $2f_1-f_h$ which is predicted from the power-series hypothesis is, in fact, generated by some other form of nonlinearity? Perhaps, the 1dB/dB growth is an artifact of method.

All previous measures of $2f_1-f_h$ were obtained with a paradigm in which it is assumed that the level of a tone adjusted in phase to cancel the perception of the subjective tones is equal in amplitude to the distortion product. Possible masking of the combination tone suggested by Lord Rayleigh (1894) and demonstrated in Experiments One and Two which might bias this interpretation had been rejected as incorrect by Goldstein.

To examine the nature of the growth of $2f_1-f_h$ and the possibility of methodological artifact, we employed two different measurement paradigms: the first was a cancellation technique similar to Goldstein's to which corrections for masking of the difference tone could be applied. The second was a binaural alternate loudness balance procedure.

The second method had the advantage that the comparison tone would not interact with the primaries thus obviating any possible basis of relative phase between stimulus components or envelope effects (Plomp, 1967).

I. METHODS

A. Cancellation Paradigm

The subject, seated in a sound-shielded room, is presented with the following stimuli monaurally:

1. A signal at frequency f_1 (1000 Hz) at predetermined sound pressure level.
2. A signal at frequency f_h (1200 Hz) at the same level as the component at f_1.
3. An interrupted signal at $f_{DP} = 2f_1 - f_h$ the amplitude and phase of which could be controlled by the listener.

The order of presentation of the primary tone levels was varied non-systematically between 40 and 80 dB SPL in 10 dB increments, each trial being conducted at a single level. The interrupted tone was fully ON for 60 msec. and OFF for 400 msec. Tone burst rise/fall times were 50 msec. A light on the subject response box pulsed ON concurrently with the tone.

The listener performed two steps: the adjustment of phase and a series of amplitude measurements. First the observer adjusted the phase of the cancelling tone (f_{DP}) in the presence of the masking primaries for minimum loudness. Subjects manipulated the phase shifter

and were instructed to bracket the setting that produced minimum loudness before reaching a final decision. No distinguishing marks were visible on the phase control knob. After establishing cancellation phase, the f_{DP} was re-introduced at a level high enough to be audible as illustrated in Figure 4.1. The subject reduced the level (L_{DP}) of the tone to the point below which she could not detect a distinct pulsed tone coincident with the cueing light (level #2). The cancellation tone amplitude was further reduced to level 3, the point at which the light and tone appeared ON alternately, after which the listener further reduced L_{DP} until the pulsing tone and light were ON concurrently (level 1). Further reduction in L_{DP} caused cessation of all pulsation at f_{DP}. The subject then allowed L_{DP} to increase, noting the level at which the tone and light were ON alternately (level 4) and finally, noting when they were ON concurrently.

Estimates of distortion product amplitude may be calculated from the above measurements with the following logic. First assume that the f_{DP} is masked. At levels 1 and 2 the vector sum of the aural distortion product (L_{ADP}) and the acoustic cancellation tone (L_{DP}) are at masked threshold as shown in Figure 4.2. At level 1, the resultant of $L_{ADP} + L_{DP}$ is also equal to masked threshold

$$MT = L_{ADP} - L_{DP_1} \qquad (4.1)$$

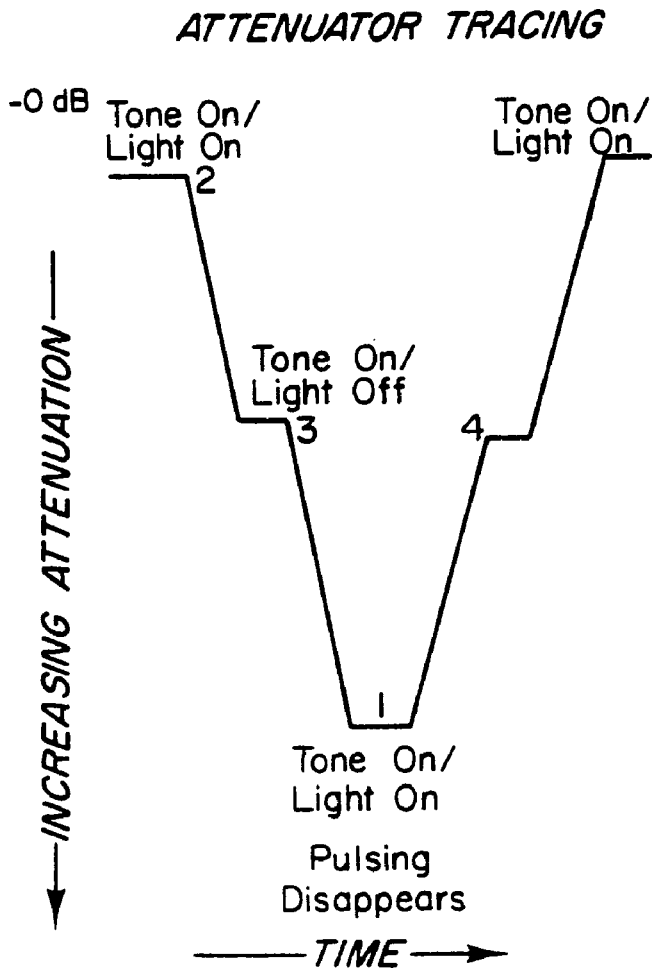


Figure 4.1: Recording attenuator tracing in the cancellation paradigm.

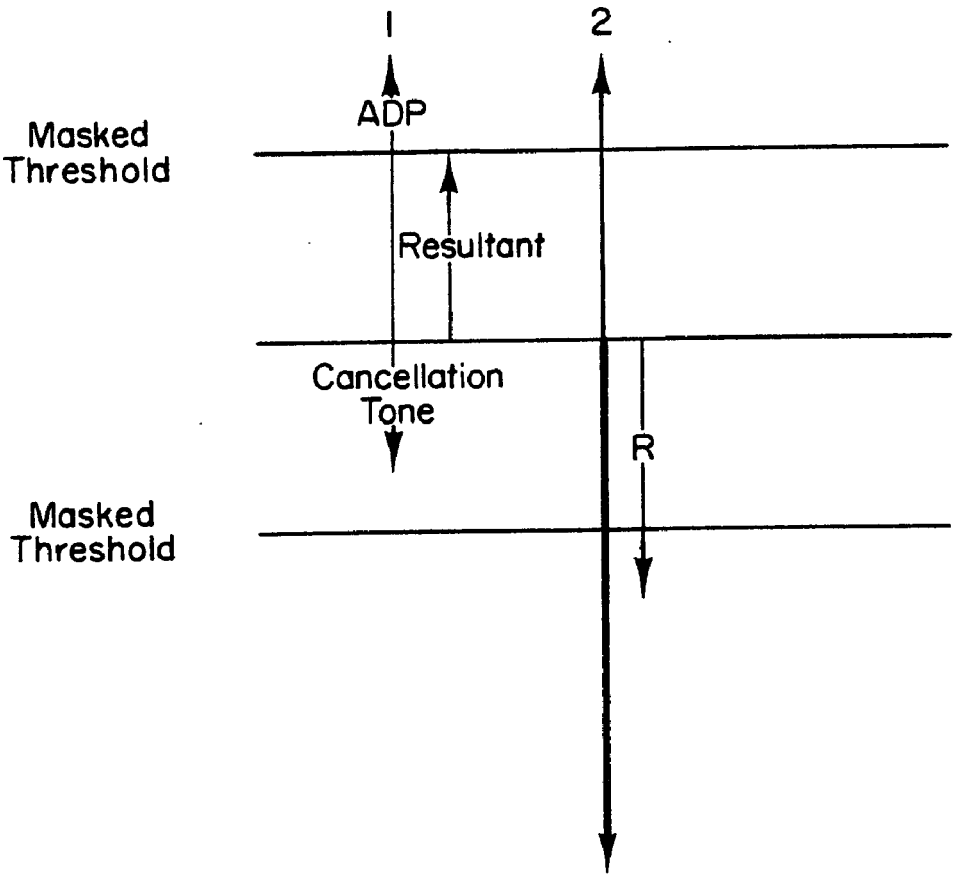


Figure 4.2: Vectorial representation of tones in the cancellation paradigm.

and at level 2 the resultant is at masked threshold and additionally is 180^{o} different from the resultant at level 1, therefore

$$-MT = L_{ADP} - L_{DP_2}. \qquad (4.2)$$

Solving equations (4.1) and (4.2) simultaneously for MT,

$$MT = (L_{DP_2} - L_{DP_1})/2. \qquad (4.3)$$

Substituting this result back into equation 4.1 the solution for the aural distortion product amplitude is

$$L_{ADP} = (L_{DP_2} + L_{DP_1})/2. \qquad (4.4)$$

B. Loudness Balance Paradigm

The loudness balance paradigm for audible subjective tone measurement has the advantage that subjective tone amplitude estimates are made in intervals containing only the two primaries in the ear generating the distortion product.

The object of the method is to equate the loudness of a subjective tone ipsilaterally and an alternately occurring pure tone at the same frequency contralaterally. This is accomplished by presenting f_{DP} binaurally and adjusting the ipsilateral acoustic tone at f_{DP} to cancellation phase; when f_{DP} is ON the distortion product in the ipsilateral ear is cancelled, but the contralateral tone at f_{DP} is audible.

The ipsilateral tone in the presence of continuous maskers of 1200 Hz and 1000 Hz was fully on 600 msec. with

50 msec. rise/fall times. The contralateral tone was ON 400 msec. with 50 msec. rise/fall times (see Figure 4.3). The effect, when the ipsilateral f_{DP} was in cancellation phase with the subjective tone was of a 400 msec. ipsilateral tone alternating with a 400 msec. contralateral tone.

The listener's task was as follows: first adjust the ipsilateral maskee (f_{DP}) for cancellation phase. This was indicated by tone ON-light OFF. Next with f_{DP} presented to both ears simultaneously, adjust the level of the contralateral tone to be equal in loudness to that of the ipsilateral tone. The second step was then repeated with an additional contralateral masker at f_1 6 dB higher in level than the ipsilateral f_1.

The loudness balance task was attempted by all subjects for primary tone levels of 50-80 dB SPL in 10 dB increments. As in the cancellation task, the order of level of presentation was varied nonsystematically.

As control for the effects of reference ear choice and order of presentation of the reference tone subjects performed an alternate binaural loudness balance at 70 dB SPL (reference ear) at both 800 Hz, the frequency of $2f_1-f_h$ and at 1000 Hz, the frequency of the lower primary. Selection of reference ear and order of presentation were balanced. At 800 Hz, amplitude differences for judgements of equal loudness for individual listeners were within 5 dB. Addition of masking as shown in

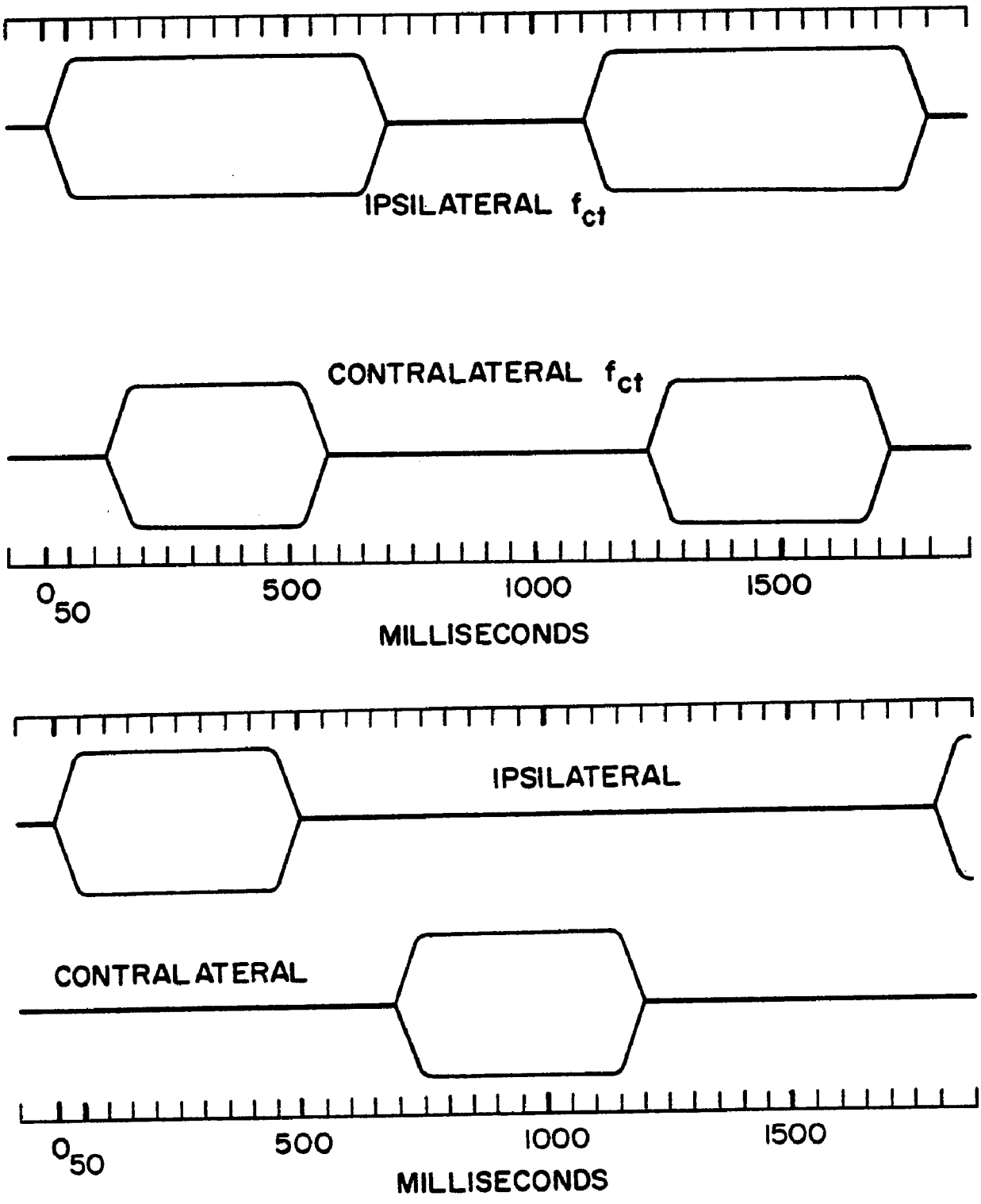


Figure 4.3 (a): Stimulus presentation sequence for loudness balance of subjective ADP ipsilaterally with physical f_{CT} contralaterally. (b) Equivalent stimulus presentation for control experiments.

Figure 4.4 had less than a 5 dB effect on the non-masked judgements; it generally reduced the difference between reference and adjusted signal levels.

II. LEVEL OF THE DIFFERENCE TONE $2f_1-f_h$

Masking corrected cancellation measures of $2f_1-f_h$ are shown in Figure 4.5.

Although all subjects reported the perception of subjective tones, only three of the listeners were able to perform the cancellation at primary levels below 60 dB SPL. Distortion product amplitudes which averaged about 25 dB below the level of the primary tone grew at rates between 0.4 and 2.4 dB/dB. The variability of growth rates is such that it is not feasible to conclude that the slope of the growth function for $2f_1-f_h$ is 1 dB/dB as has been reported (Goldstein, 1967; Hall, 1972a). The binomial test indicates that the number of slopes greater than 1 dB/dB is at the rate expected by chance. There is, however, more than a 10% probability that if the slope of the function were 2.4 dB/dB that five of the six subjects would, by chance, manifest lower slopes.

For comparison with the cancellation estimates corrected for masking effects in Figure 4.5, the data of Figure 4.6 are cancellation estimates derived from the same measurements as the data of Figure 4.5 using the point of best cancellation, (the point at which the TONE-ON/LIGHT-OFF perception was strongest) as the measure of L_{ADP}. This is the type of measurement reported by Goldstein. In this

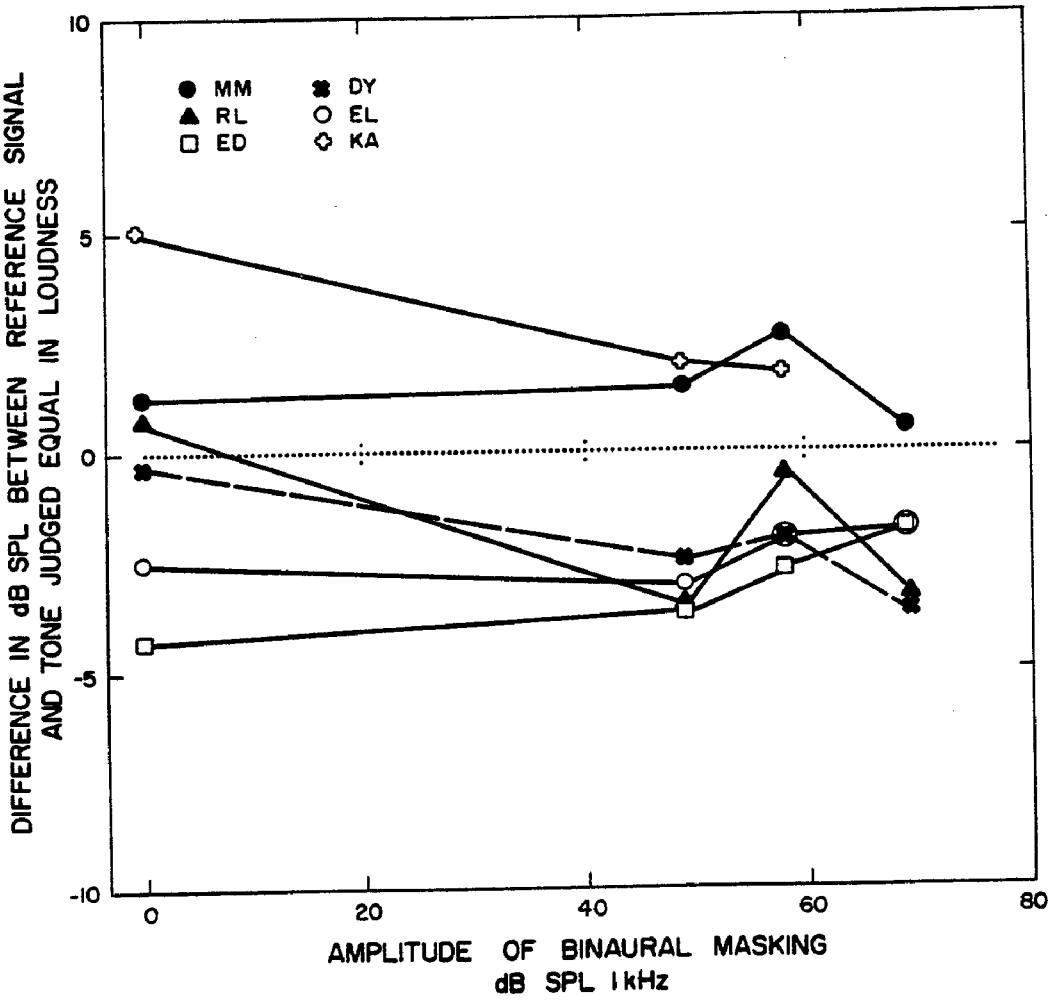


Figure 4.4: Differences between signal amplitudes judged equal in loudness under conditions of diotic masking. Maskee is 800 Hz, 70 dB SPL, masking signal is 1000 Hz at amplitude indicated on abscissa.

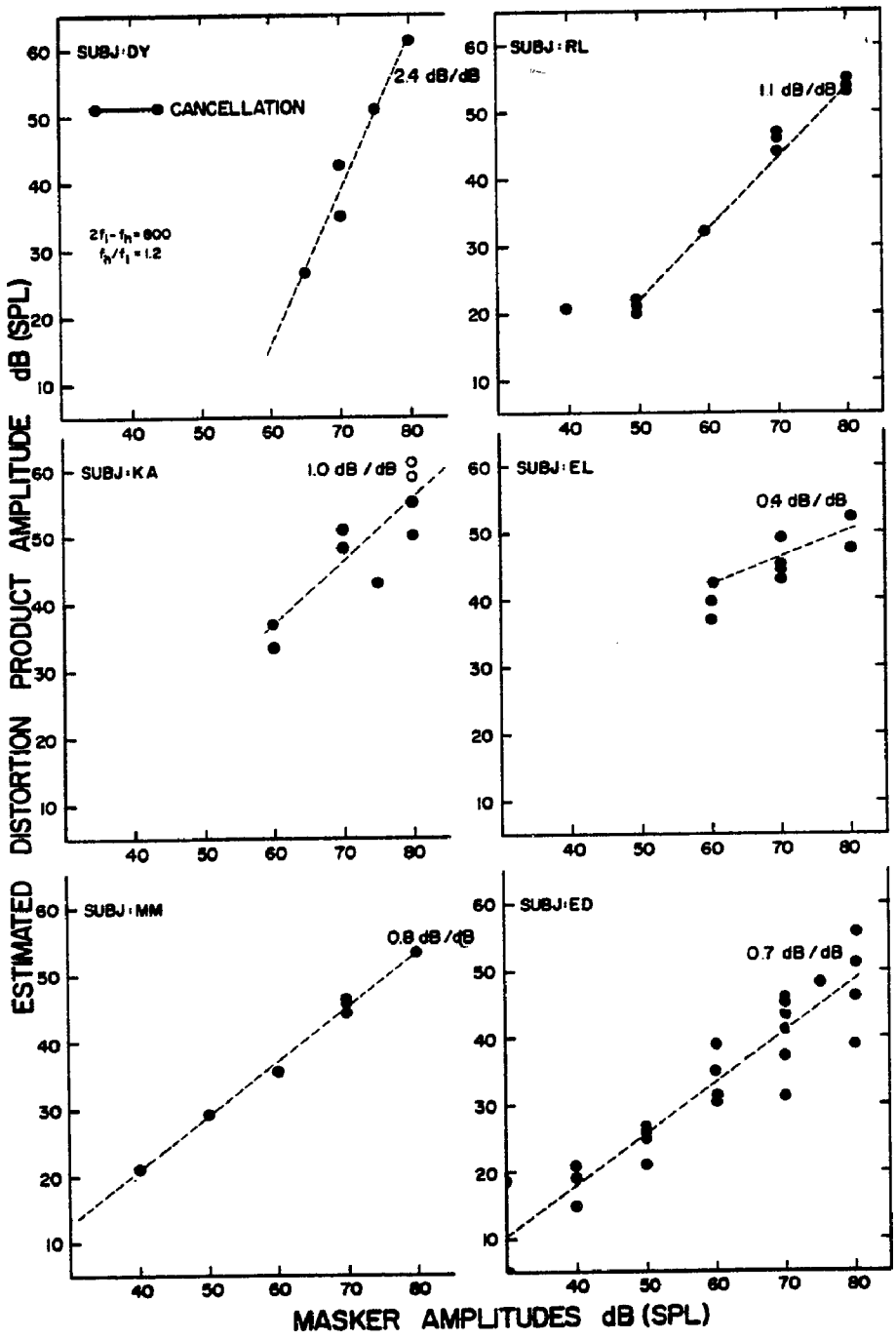


Figure 4.5: Measures of the audible difference tone $2f_1-f_h$; $f_h/f_1 = 1.2$, $f_1 = 1000$ Hz.

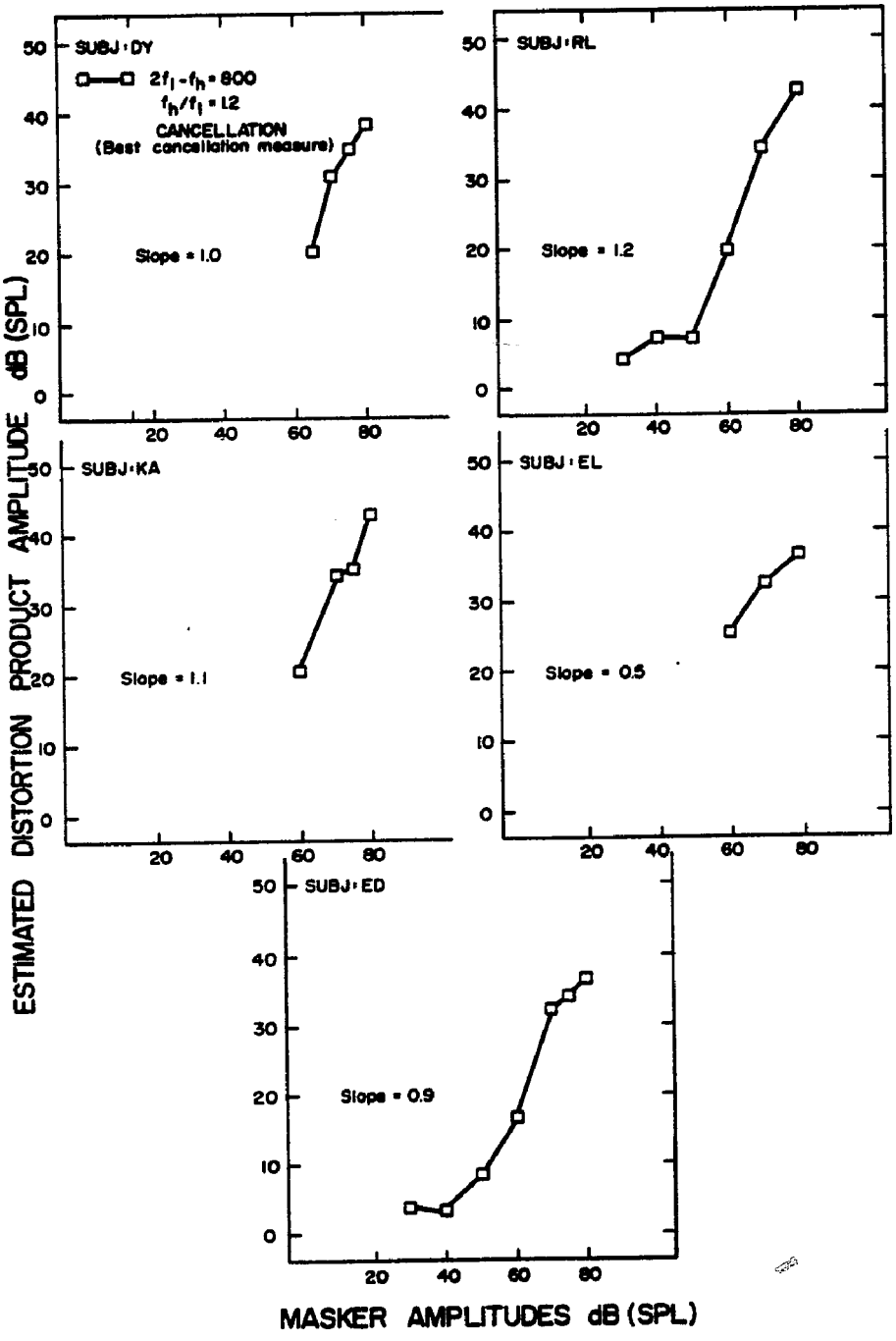


Figure 4.6: Cancellation estimates of $2f_1-f_h$ not corrected for masking.

case, the slopes of the $2f_1-f_h$ growth functions in the region of increasing distortion product level including that of DY approximate 1 dB/dB (DY = 1.0, RL = 1.2, KA = 1.1, EL = 0.5, ED = 0.9).

Considering that correction of cancellation measures for masking effects results in increased variability of growth rates, it appears that the 1dB/dB growth reported in the literature may be an artifact contributed to by masking from the primaries.

Loudness balance measures of $2f_1-f_h$ tended to support the idea that 1dB/dB growth of $2f_1-f_h$ is artifactual. Figure 4.7 shows the amplitude estimates obtained for five of the listeners.

With the exception of the data for EL, distortion product amplitudes estimated from loudness balance measures grow at 3 dB/dB or less. The data of EL are similar to behavior observed for the inaudible distortion products - an asymptotic minimum distortion product level followed by an increase with primary tone intensity of approximately 3dB/dB. The one other abberant point (ED, 80 dB SPL) shows a sharp decrease (dB/dB) in the estimated distortion product from 70 to 80 dB. At this time it is not possible to explain the behavior of ED between 70 and 80 dB SPL nor can any additional insight be made with reference to the slope of the growth function of $2f_1-f_h$.

III. DISCUSSION

The major feature of the growth of the difference tone

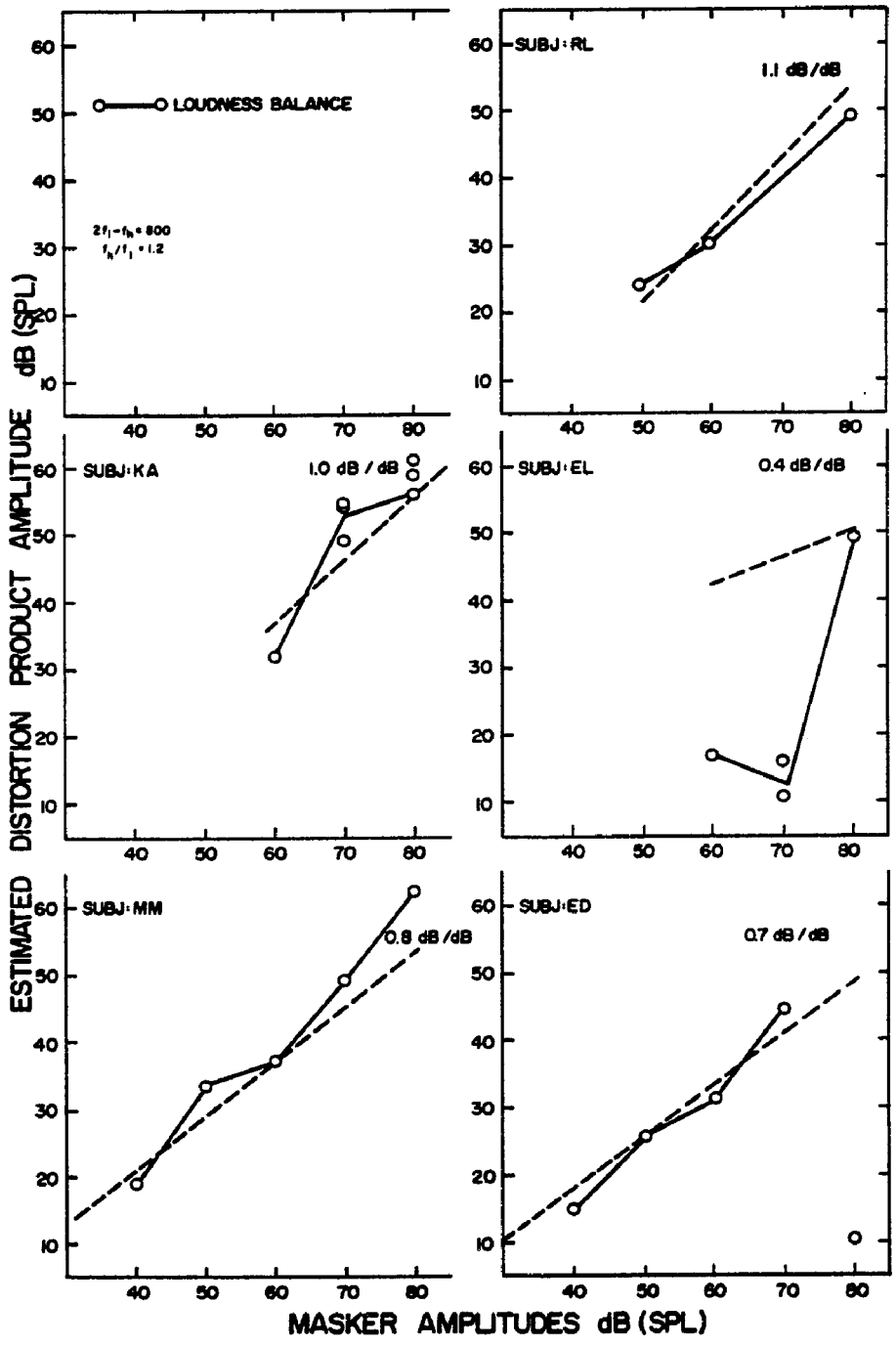


Figure 4.7: Measures of the audible difference tone $2f_1-f_h$; $f_h/f_1 = 1.2$, $f_1 = 1000$ Hz using loudness balance paradigm. The broken line is the fit of the "corrected cancellation" data for comparison.

$2f_1-f_h$ in this experiment is the variability of the growth rate between individuals and experimental methods. It is significant that the same cancellation data from a single individual analyzed in two similar ways (e.g. corrected and uncorrected for masking) can yield different rates of growth of the distortion product. We infer from this that the true nature of audible difference tone generation is obscured by methodology. We cannot conclude from these measurements that the growth of the difference tone $2f_1-f_h$ is 1dB/dB, greater than 1dB/dB, or less than 1dB/dB.

CHAPTER FIVE

THREE MODELS OF AUDITORY NONLINEARITY

The demonstration of summation tones in Chapter Three and of combination tone masking in that chapter and in Chapter Two has provided the necessary evidence for a complete comparison of the three models of auditory nonlinearity. The relationship of the results of the three experiments to Helmholtz' original explication of auditory amplitude distortion, to Fletcher's simplification of Helmholtz' theory, and to Goldstein's essential nonlinearity will be considered in order.

I. HELMHOLTZ' THEORY

The mechanism of combination tone generation proposed by Helmholtz produces subjective tones proportional to their frequency. The amplitude of the quadratic combination tones $f_h \pm f_1$ are given by Helmholtz (p. 412) as

$$L_{f_h \pm f_1} = \frac{uv}{m(f_1 \pm f_h)^2 - a}.$$

In this case $u = \frac{P_1}{mf_1^2 - a}$,

$$v = \frac{P_h}{mf_h^2 - a},$$

P_h and P_1 are the sound pressure levels of each primary. If the frequency of the lower primary is held constant, the level of the combination tones will be a complex function

of the frequency of the upper primary, assuming that the level of each tone is fixed. The precise nature of the relationship depends on the values of "m" and "a". As the frequency separation between primaries increases, the level of the sum tone must decrease and the level of the difference tone may increase. The situation is similar for the cubic distortion products. These combination tones increase with masker intensity at an exponential rate determined by the order of the term of the power-series from which they are derived. For this reason, combination tones of the first order ($f_h \pm f_l$) and aural harmonics of the same order (2f) should each have different amplitudes but similar rates of growth. Without knowledge of the specific subjective tone frequencies the level of one tone cannot be predicted from another. As was shown in Chapters Two and Three, however, combination tone levels may be predicted accurately from aural harmonic estimates. This suggests that the nonlinear mechanism producing the distortion product is sensitive only to the level of the stimulus at its input and not its frequency.

The differences between the levels of each tone of the pair $2f_h \pm f_l$ further refute the precise accuracy of Helmholtz' theory. For the subjective tones $2f_h \pm f_l$ with $f_h/f_l = 1.6$ and f_l = 1KHz, the difference between the two tones should be on the order of 3200 Hz/2200 Hz or 1.45 times, the difference tone being 3.25 dB higher than the sum tone. If the factors "m" and "a" are dominant, however, then the

subjective tone amplitudes will be similar; but in no case can this mechanism account for the 25 dB difference found between $2f_h+f_1$ and $2f_h-f_1$.

In summary, the results of the first two experiments invalidate certain details of Helmholtz' model, primarily the dependence of amplitude on distortion product frequency. The close agreement between values of f_h+f_1 predicted from aural harmonic estimates and the subjective tone amplitudes measured with a two-tone masker are also inconsistent with the sensitivity to distortion product frequency inherent in this model. Agreement found between amplitudes of $2f_1+f_h$ and $2f_h+f_1$ similarly is inconsistent with this aspect of the model. The observation that in some instances the level of $2f_1+f_h$ increases with frequency separation is further evidence against a frequency selective nonlinear model as proposed by Helmholtz.

The major supported aspect of Helmholtz' model is production of subjective tones by an amplitude nonlinearity and accerlerated distortion product growth rate determined by the order of the term of the power-series nonlinearity by which it is generated. This is the essence of Fletcher's modification of the Helmholtz model.

II. FLETCHER'S MODEL

For a power-series nonlinearity model of the form $X = a_0+a_1p+a_2p^2+a_3p^3+\ldots+a_np^n$ the aural distortion product is proportional to the order of the term from which it derives, i.e., the level of a subjective tone generated by

the second order term a_2p^2 is in proportion to the square of the stimulus amplitude,that derived from a_3p^3 is proportional to the cube of the stimulus level.

Substitution of $p_1 \sin \omega_1 t+p_2 \sin (\omega_2 t+\phi)$ into the power-series will show, on simplification of the result, that the third aural harmonic is 9.5 dB lower than a third order combination tone generated by the same level tone; the second harmonic is 6 dB lower than the second order combination tone. All combination tones of equal order generated by equally effective stimuli to the nonlinearity should be of equal amplitude. For equally effective stimuli, the level of the combination tone produced is frequency independent.

In general, the inaudible combination tones exhibit the accelerated growth predicted from this model. Variations in absolute levels of combination tones at different frequency are probably the result of frequency-dependent recoupling of the distortion product at its proper place after it is generated (cf. Smoorenburg, 1972b). This stage of processing follows the nonlinearity and is not directly related to the nonlinear amplitude characteristic. The fact that combination tone levels are predictable from aural harmonic estimates suggests that both are produced by the same nonlinearity and, further, that this nonlinearity is independent of the frequency of the inputs.

In general, the results of the preceding experiments support the existence of an aural nonlinearity producing

distortion products which increase at rates faster than the primary tones and which are masked by their generating tones.

III. ESSENTIAL NONLINEARITY

Neither of the preceding two models explain the non-accelerated growth found for distortion products lower in frequency than the primaries. These have been explained in other reports (Goldstein, 1967) by postulating a level-normalized nonlinear mechanism in conjunction with a spatial filter which effectively eliminates the subjective tones higher in frequency than the primaries.

The essential nonlinearity was proposed in part because of the inability to explain the absence of summation tones and in part because "the studies of cubic CT-level behavior showed that its relative distortion is nearly independent of stimulus level but is sharply dependent upon stimulus-frequency interval." Each of these observations is at variance with our observations for the distortion products above the frequency of the primaries. We have not shown, as Goldstein and as Zwicker did for f_h-f_1, that the relative distortion product amplitude is dependent on intensity. Furthermore, on our six subject population we have been unable to support a frequency selective nonlinearity. The last difference between Goldstein's results and ours, also a reason for his essential nonlinearity hypothesis, was that we have demonstrated masking of distortion products at $2f_h-f_1$ where he claimed it did not exist (p. 686).

Differences in frequency separation effects found in this study and in Goldstein's could be the result of the different distortion products examined. Goldstein made his measurements at $2f_1-f_h$, the cubic difference tone whereas ours were made at $2f_1+f_h$, the cubic summation tone. If the combination tones are produced by a single nonlinearity of the power-series type or an essential nonlinearity, however, both distortion products should behave similarly. At this point no complete explanation can be offered for the differences in frequency selectivity. Differences in our conclusions regarding inaudibility by masking, however, may be explained on the basis of his methods. Goldstein concludes that inaudibility by masking at $2f_1-f_h$, "...is proven unacceptable...because CT's occur at sufficiently low levels to avoid excessive high-frequency masking by the primary stimulus." At $2f_h-f_1$ he concedes that masking exists but, on the basis of the level of $2f_1-f_h$, it "is subtle; but here too, masking does not account for the absence of the CT." If his estimates of the level of $2f_1-f_h$ are in error, so may be his arguments against inaudibility by masking.

Listeners in his study raised the cancellation tone until, in the presence of masking by the primaries, it became audible. Cancellation level was approached only from one direction which may bias the resultant value determined for $2f_1-f_h$ level. Goldstein cites as evidence against the masking hypothesis the fact that $2f_1-f_h$ was audible at 50 dB SPL and that the masked threshold at the third harmonic "was

found to be insignificantly different from the quiet threshold which,of course,was much lower than the level of CT $2f_1-f_h$." At $2f_h-f_1$, "Maskee phase had little effect (less than 2 dB) upon threshold...". From this he concludes that masking at these frequencies is inconsequential.

In contrast, our results indicating that, in addition to the existence of masking at these frequencies, small phase effects may indicate the presence of real distortion products. Data (at f_h+f_1) from this study with a phase effect of less than 2 dB at 50 dB SPL primaries, when analyzed, yielded a masked threshold of less than 4 dB SPL and a reliable distortion product of -10 dB SPL. Similar data at $2f_1+f_h$ yielded distortion estimates about -20 dB SPL.

Further comparison between Goldstein's results (subj. JLG) and ours from the cancellation paradigm show his distortion product estimates at 50 dB SPL to be about 20 dB higher than ours (25 dB SPL, JLV vs. 5 dB SPL average of our subjects using the masking-corrected analysis). In this study with 70 dB SPL primaries, the cubic difference tone $2f_1-f_h$ estimated from points of best cancellation ranges between 15-25 dB SPL. If the power-series hypothesis is valid, the companion cubic summation tones would be predicted at the same level. These were measured at 8-20 dB SPL by tone-on-tone masking - a discrepancy of only 5-7 dB.

Masking by a single 70 dB, 1 KHz primary at about 1.5 octaves above its frequency is estimated from previous

masking data (Ehmer, 1959) at about 30 dB-10dB above the highest measured cubic summation tone. Adding the effect of masking by the lower primary would increase the total masking at the distortion product frequency. Masking, therefore, would be expected to cause the sum tone to remain inaudible. Furthermore, if masking grows at 2dB/dB and cubic sum tones at 3dB/dB, as the primary level is decreased, the distortion product will decrease faster than the masking; the result being that at low levels masking is more of a factor than at high.

Distortion products may be generated at levels far below even quiet thresholds (Clack, 1975). Masking is a viable explanation of the inaudibility of aural distortion products; a post-nonlinearity filter need not be invoked to remove high frequency distortion products.

The precise nature of the growth of $2f_1-f_h$ and of f_h-f_1 is still uncertain. The observation of $2f_1-f_h$ growth of 2.4dB/dB for DY, the 1dB/dB growth of f_h-f_1 for one listener, and the differences in growth rates when cancellation measures are compared with and without masking correction each contribute to this uncertainty. Because of this uncertainty, the need for a separate mechanism to explain the characteristics of $2f_1-f_h$ and f_h-f_1 is not clear.

IV. CONCLUSIONS

The results of these experiments suggest that a frequency independent nonlinear mechanism such as proposed by Fletcher may have the greatest descriptive utility in ex-

plaining the observed behavior of aural combination tones. This model must be modified with selective filtering at both the input and the output to reflect the differences in interaction between primary tones of different frequency and to correct for the recoupling of the distortion product back into the perceptual pathway.

If the growth of $2f_l-f_h$ is eventually confirmed as 1dB/dB, this would be inconsistent with the power-series nonlinearity model. Two possible explanations exist for this growth rate other than its origin in a second nonlinearity. The first is that suprathreshold distortion products, because of their audibility, are subject to additional levels of processing which somehow result in a 1dB/dB growth characteristic. This possibility is currently being examined in a study of $2f_l \pm f_h$ in which the difference product is artificially rendered inaudible with narrow band masking. The other possibility may be found in the interaction between a subjective tone perceived analytically (a nonlinear distortion product) and a subjective tone at the same frequency perceived synthetically (a subharmonic such as investigated by Houtgast, 1976).

CHAPTER SIX

SUMMARY

Measurements were made of three classes of aural distortion product: harmonics, summation tones, and difference tones. Two major results of these measurements are: (1) the existence of aural summation tones has been demonstrated; (2) each distortion product is a subset of combination tones predictable from a power-series amplitude nonlinearity, and (3) the amplitude each class of distortion product may be predicted from the other classes.

The rate of growth and, within certain limitations, the absolute levels of combination tones higher in frequency than the primary tones which generate them are predictable from a power-series nonlinearity of the form $X = a_0+a_1p+a_2p^2+a_3p^3+\ldots+a_np^n$.

Comparison of estimates of $2f_1+f_h$ with f_1 = 1000 Hz and f_h = 1200 Hz or 1600 Hz indicate that the nonlinearity is independent of frequency. That is, the level of distortion product produced by the nonlinear amplitude characteristic is a function only of the stimulus input level.

The nonlinearity is probably preceded and followed by frequency dependent filters.

Measurements of the audible difference tone $2f_1-f_h$ with three methods were highly variable and did not conclusively demonstrate a 1dB/dB growth characteristic. The first-order difference tone f_h-f_1 also increased at 1dB/dB.

We conclude that the amplitude nonlinearity is best described by a power-series with filtering preceding and following the

site of the nonlinearity.

The direction of future related studies relate to the closer delineation of the nature of the nonlinearity and in the application to other areas quantification of the nonlinearity. Specifically, the growth of $2f_1-f_h$ and f_h-f_1 remain to be explained. It is possible to test the hypothesis that the differences in growth are attributable to audibility simply by masking $2f_1-f_h$ and estimating its amplitude from tone-on-tone threshold data. If the growth is found to be 1dB/dB in this case, further studies will be undertaken to examine the possibility of interaction between synthetically and analytically perceived pitches.

In the area of application, I plan to use these methods to reexamine an old hypothesis that distortion as an indicator of "aural overload" may be used to predict susceptibility to auditory fatigue (Lawrence and Blanchard, 1954).

APPENDIX

ESTIMATE ACCEPTABILITY (ACCEPTANCE CRITERION RATIO)

The observation that aural harmonic estimates increase at 2dB/dB or at 3dB/dB at primary levels above 50-60 dB SPL and at lower rates below this level raises an important question for the determination of the mechanism of nonlinearity.

The question is, "Is this the result of the distortion products (or the phase effects) arising from different mechanisms as a function of stimulus level (Tonndorf, 1970; Craig and Jeffress, 1962), or is it an indicator of noise dependent limits to the lowest level of a distortion product which can be measured (Dallos, 1973)?" If the latter explanation is correct, increases in distortion product amplitude should result in increases of signal-to-noise ratios which then will yield more reliable estimates of subjective tone amplitudes.

Previous criteria for the determining validity of distortion estimates derived from monaural phase effect measures required that the phase effect be at least 0.5 dB and that the sum of squared residuals (SSR) be less than 25 dB^2 (Clack, et. al., 1968 unpublished). These criteria have the disadvantage that the absolute value of the SSR above which the estimates are considered invalid is unrelated to the overall amplitude of the monaural phase effect. Furthermore, at low masker levels possibly valid estimates may be rejected simply on the basis of small amplitude monaural phase effects. As part of this study, an alternative criterion of estimate va-

lidity was developed.

This statistic, the Acceptance Criterion Ratio (ACR), is independent of masker amplitudes and frequencies and is applicable over all magnitudes of the phase amplitude effect. The ratio is calculated by first computing the sum of the squared residuals (SSR) between the fitted function and the data. The square root of the SSR is divided by the estimated distortion product amplitude (ACR $=\sqrt{SSR/L}$, where L is the distortion product amplitude) to yield the ACR. As in the analysis procedure in the preceding section, all calculations are in dynes/cm^2.

This statistic has the advantage that the measure of dispersion about the fitted curve (the square-root of the sum of squared deviations) is normalized with regard to the amplitude of the estimated distortion product. For an intuitive understanding of the operation of the ACR, consider the following example. If all the thresholds except one lay on the fitted equation and that one point was in error by an amount equal to the distortion product amplitude (L), then the ACR would equal one (ACR $=\sqrt{L^2/L} = 1$. A perfect fit would yield an ACR equal to zero. At the opposite extreme six data points of noise equivalent in amplitude to the distortion product and having no sinusoidal component would yield an ACR $= \sqrt{6L^2/L} = 2.4$.

A ratio of 1.3 was chosen as a cutoff above which subjective tone estimates based on tone-on-tone masking data would be considered invalid. For lack of concrete bases on which to base this decision, the cutoff point was determined by consideration of data from previous studies which were thought to

be well fit by the descriptive function.

As such, the ACR represents a quantitative basis on which estimates are accepted or rejected; the value of the ACR chosen as cut-off in this study is not based on objective investigation of the manner in which the ACR varies with proportionate changes in signal/noise ratio.

BIBLIOGRAPHY

Bekesy, G. von: Experiments in Hearing. McGraw-Hill, New York, 1960.

Boer, E. de: A Note on Phase Distortion and Hearing. Acustica, 11:182-184, 1961.

Chapin, E. K., and Firestone, F. A.: The Influence of Phase on Tone Quality and Loudness; the Interference of Subjective Harmonics. J. Acoust. Soc. Am. 5:173-180, 1934.

Clack, T. D.: Aural Harmonics: The Masking of a 2000-Hz Tone by a Sufficient 1000-Hz Fundamental. J. Acoust. Soc. Am. 42:751-758, 1967.

Clack, T. D.: Aural Harmonics: Preliminary Time - Intensity Relationships Using the Tone-On-Tone Masking Technique. J. Acoust. Soc. Am. 43:283-288, 1968.

Clack, T. D.: Some Influences of Subjective Tones in Monaural Tone-On-Tone Masking. J. Acoust. Soc. Am. 57:172-180, 1975.

Clack, T. D., Erdreich, J. and Knighton, R. W.: Aural Harmonics: The Monaural Phase Effects at 1500 Hz, 2000 Hz, and 2500 Hz Observed in Tone-On-Tone Masking When f_1 = 1000 Hz. J. Acoust. Soc. Am. 52:536-541, 1972.

Clack, T. D., Knighton, R. W., and Clack, A. M.: Aural Harmonics: Estimates of Aural Harmonic Amplitudes from Tone-On-Tone Masking Measurements. Unpublished Manuscript, 1968.

Dallos, P.: The Auditory Periphery. Biophysics and Physiology. Academic Press, New York, 1973.

Egan, J. P., and Hake, H. W.: On the Masking Pattern of a Simple Auditory Stimulus. J. Acoust. Soc. Am. 22:622-630, 1950.

Egan, J. P., and Klumpp, R. G.: The Error Due to Masking in the Measurement of Aural Harmonics by the Method of Best Beats. J. Acoust. Soc. Am. 23:275-286, 1951.

Ehmer, R. H.: Masking by Tones vs. Noise Bands. J. Acoust. Soc. Am. 31:1253-1256, 1959.

Erdreich, J., and Clack, T. D.: An Algorithm for Analysis of Phase Effects: Application to Monaural Distortion Product Estimation. J. Acoust. Soc. Am. 52:1124-1126, 1972.

Fletcher, H.: Speech and Hearing. Van Nostrand and Co., New York, 1929.

Goldstein, J. L.: Auditory Nonlinearity. J. Acoust. Soc. Am. 41:676, 1967.

Goldstein, J. L., and Kiang, N.Y.S.: Neural Correlates of the Aural Combination Tone $2f_1-f_2$. Proc. of the IEEE. 56:981, 1968.

Hall, J. L.: Monaural Phase Effect: Cancellation and Reinforcement of Distortion Products f_2-f_1 and $2f_1-f_2$. J. Acoust. Soc. Am. 51:1872-1881, 1972a.

Hall, J. L.: Auditory Distortion Products f_2-f_1 and $2f_1-f_2$, J. Acoust. Soc. Am. 51:1863-1871, 1972.

Hällström, G. G.: Von den Combinationstönen: Annalen der Physik und Chemie (Leipzig) 24:438, 1832.

Helmholtz, H. L. F.: Die Lehre von den Tonempfindungen, 1855. English translation On the Sensations of Tone, by A. J. Ellis. Dover Publications, New York, 1954.

Houtgast, T.: Subharmonic Pitches of a Pure Tone at Low S/N Ratio. J. Acoust. Soc. Am. 60:405, 1976.

Jones, A. T.: The Discovery of Difference Tones. American Physics Teacher. 3:49-51, 1935.

Lawrence, M. and Yantis, P. A.: Onset and Growth of Aural Harmonics in the Overloaded Ear. J. Acoust. Soc. Am. 28: 855-858, 1956.

Lewis, D., and Larsen, M. J.: The Cancellation Reinforcement and Measurement of Subjective Tones. Nat'l. Acad. Sci. Proc. 23:415, 1937.

Plomp, R.: Beats of Mistuned Consonances. J. Acoust. Soc. Am. 42:462, 1967.

Lord Rayleigh (J. W. Strutt): The Theory of Sound, 1894, Reprinted 1945, Dover Publications, New York.

Siegel, S.: Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York, 1956.

Small, A. M.: Periodicity Pitch, in Foundations of Modern Auditory Theory. Vol. 1., Tobias, J. V. ed. Academic Press, New York, 1970.

Smoorenburg, G. F.: Audibility of Combination Tones. J. Acoust. Soc. Am. 52:603, 1972a.

Smoorenburg, G. F.: Combination Tones and Their Origins. J. Acoust. Soc. Am. 52:615, 1972b.

Trimmer, J. D., and Firestone, F. A.: An Investigation of Subjective Tones by Means of the Steady Tone Phase Effect. J. Acoust. Soc. Am. 9:24, 1937.

Wegel, R. L., and Lane, C. E.: Auditory Masking of One Pure Tone by Another and Its Probable Relation to the Dynamics of the Inner Ear. Phys. Review. 23:266, 1924.

Young, T.: Outlines of Experiments and Inquiries Respecting Sound and Light. Phil. Trans. Royal Soc. London. 90: 604, 1800.

Zwicker, E.: Der Ungewönhnliche Amplitudengang der Nichtlinearen Verzerrungen des Ohres. Acustica 5:67-74, 1955.

Zwicker, E.: Influence of a Complex Masker's Time Structure on Masking. Acustica 34:138-146, 1976.