#1

PERCEPTION OF MUSICAL INTERVALS: EVIDENCE FOR THE CENTRAL ORIGIN OF THE PITCH OF COMPLEX TONES

ADRIANUS J. M. HOUTSMA
JULIUS L. GOLDSTEIN

LOAN COPY

TECHNICAL REPORT 484

OCTOBER 1, 1971

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS
CAMBRIDGE, MASSACHUSETTS 02139

The Research Laboratory of Electronics is an interdepartmental laboratory in which faculty members and graduate students from numerous academic departments conduct research.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, by the JOINT SERVICES ELECTRONICS PROGRAMS (U. S. Army, U. S. Navy, and U.S. Air Force) under Contract No. DAAB07-71-C-0300, and by the National Institutes of Health (Grant 5 PO1 GM14940-05).

Requestors having DOD contracts or grants should apply for copies of technical reports to the Defense Documentation Center, Cameron Station, Alexandria, Virginia 22314; all others should apply to the Clearinghouse for Federal Scientific and Technical Information, Sills Building, 5285 Port Royal Road, Springfield, Virginia 22151.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS

Technical Report 484 October 1, 1971

PERCEPTION OF MUSICAL INTERVALS: EVIDENCE FOR THE CENTRAL ORIGIN OF COMPLEX TONES

Adrianus J. M. Houtsma and Julius L. Goldstein

This report is based on a thesis by A. J. M. Houtsma submitted to the Department of Electrical Engineering at the Massachusetts Institute of Technology, June 1971, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

(Revised manuscript received June 25, 1971)

Abstract

Melodies can be recognized in music, regardless of the instrument on which they are played. This is true even when the musical sounds have no acoustical energy at the fundamental frequency. This phenomenon has been investigated qualitatively and quantitatively through a series of experiments in which subjects were asked to identify melodies and simple musical intervals. Each musical note was played by a complex tone comprising successive upper harmonics with randomly chosen lower harmonic number. Melodies and intervals played with such sounds consisting of only two partials of low harmonic number could be identified perfectly both when the complex tones were presented monotically (both partials to one ear) and when the partials were distributed dichotically (one partial to each ear). Control experiments showed that neither difference tone nor transformations based on frequency difference per se explain these phenomena. Percent correct identification decreased both with increasing fundamental frequency and average harmonic number. Performance is essentially the same for monotic and dichotic stimulus paradigms, except for differences which were shown to be accounted for by aural combination tones. Moreover, identification performance is essentially random when the harmonic numbers of the stimulus tones or audible aural combination tones are sufficiently high (greater than approximately 10) so that they cannot be resolved behaviorally in monaural experiments. These findings suggest that sensations of "musical pitch of the fundamental" in complex tones are mediated centrally by neural signals derived from those stimulus partials that are tonotopically resolved, rather than being mediated by neural transformations of those upper partials which the peripheral auditory system fails to resolve.

Similar experiments with complex tones containing up to six successive partials yielded results that are basically similar to those with the two dichotic partials. Thus the basic properties of the central processor of musical pitch can be studied under stimulus conditions that are free from the confounding and irrelevant effects of cochlear nonlinear interactions among partials.

Phase effects were studied for monotic and dichotic tone complexes and the results suggest that the central pitch processor is insensitive to the relative phases of its separate inputs. Some experiments were performed on melody perception using inharmonic two-tone stimuli and the results are consistent with earlier studies of musical pitch.

A preliminary attempt was made to reduce the experimental data to a single parameter of a statistical decision model. Comparisons were made between the interval identification data and optimum processor predictions based on frequency discrimination data for simple tones.

TABLE OF CONTENTS

CONTENTS

I. INTRODUCTION

The human ear has been the object of numerous theoretical and experimental studies for centuries, and almost invariably each conceptual or experimental advance has put new obstacles in the way of understanding what was once thought to be a simple system. A survey of psychoacoustical literature of the past hundred years shows clearly how extremely sophisticated the auditory system is, and how new experimental facts appear to render the development of a complete theory of hearing more remote.

In this report we are occupied with just one of the many amazing operations that the auditory system can perform. Through many ages man has developed an activity to please himself through his sense of hearing, namely music. The human ear has the remarkable ability to track several melodies (see Appendix III for definitions) simultaneously when listening to a sequence of orchestral sounds. This ability was a basic essential of all Western music at least up to the Baroque period. A simple case of this "polyphonic music" is monophonic music, which is music comprising just one melody. A considerable part of the Western musical tradition, and almost all of the Oriental tradition, consists in this kind of music. When we consider the conditions under which the human ear can track a simple melody, we realize that the ear is performing anything but a simple operation. Over the years an internationally agreed upon note scale has been developed which characterizes musical sounds solely by their fundamental frequencies. A melody is a sequence of such notes. When such a note sequence is transformed into a sound sequence according to the rule just mentioned, the result is a sensation of the original melody; that is, the original sequence of notes can be retrieved from a listener's sensation in one form or other. The amazing thing, however, is that the melody sensation seems to be invariant over a large class of note-to-sound transformations. Melodies can be played on instruments that have quite different frequency transfer characteristics, and still be recognized. Some instruments produce sounds with a rich spectrum comprising the fundamental and many partials (stringed instruments), Some have little else than the bare fundamental (flute in high register), others generate spectra having a strong "formant" region with the result that the lower notes usually lack the lower partials, including the fundamental (oboe). It can also be observed that melodies played on any given insturment are still easily perceived after the sound has been passed through a formant filter having a bandwidth much narrower than the spectra of the original sounds. This is evident all the time when one listens to music from an inexpensive transistor radio or over a telephone line. In this study we shall attempt to find out through a series of psychophysical experiments how the auditory system retrieves and encodes the information of the fundamental frequency from a series of musical sounds.

The notion that musical sounds are sounds whose waveforms have a temporal regularity such that their value on the note scale is determined by the temporal repetition rate goes back to the seventeenth century (Mersenne,[1] 1636; Galileo,[2] 1638). How the ear tracks the successive temporal periodicities in a sequence of sounds was first seriously raised as a scientific issue, however, when Seebeck[3] published some experimental observations, in 1841. Using a siren, he was able to generate signals with ambiguous repetition rates, such as a pulse train having interpulse periods equal to t_1, t_2, t_1, t_2, etc. and t_1, t_2, t_3, t_1, t_2, t_3, etc. He observed that the ear often perceives such a sequence of irregular pulses as sequences of regular pulses with repetition periods of $t_1 + t_2$ and $t_1 + t_2 + t_3$, respectively. Only when the differences between t_1, t_2, and t_3 become very small will the ear perceive the pulse trains as if they were regular pulse trains with periods of $(t_1+t_2)/2$ and $(t_1+t_2+t_3)/3$, respectively.

Ohm,[4] in 1843, disagreed with Seebeck's notion that the ear tracks periodicities in acoustic waveforms. He stated that in order to hear a tone corresponding to the frequency f, the waveform must contain a component $A \sin(2\pi f t + \theta)$, a statement which became famous as Ohm's Acoustic law.[5] He applied Fourier's theorem to the acoustic signals used by Seebeck and showed that his theory predicted precisely what Seebeck had observed. In the discussion that followed, Seebeck[6-8] argued that Ohm's theory did not explain his observations quantitatively, since he perceived the fundamental frequency much stronger than would be predicted by Fourier's theorem, and therefore that the "higher harmonics, which share a common period, somehow add to the sensation of the fundamental." Ohm argued that Seebeck's observation were not based on physical facts but on an acoustical illusion. Seebeck then pointed out that the issue was how periodic sounds are perceived and not how they can be described physically, and that such an issue can only be decided by the ear.

Helmholtz,[9] in 1863, incorporated Ohm's law in a psychophysical theory of hearing which can be described in four parts:

1. The ear performs a Fourier analysis with limited resolving power. It analyzes a complex sound "into precisely the same constituents as are found by sympathetic resonance, that is, into simple tones, according to Ohm's definition of this conception."

2. Each resonator excites a corresponding nerve fiber, which causes a specific tonal sensation.

3. The sound-transmission process in the ear is characterized by nonlinear distortion at high sound intensities.

4. Stimulus interference produces sensation interference. Consonance is determined by the absence of beats that arise from limited spectral resolution.

The third hypothesis was necessary to explain earlier observation by Tartini,[10] Sorge,[11] Romieu,[12] and others that when the ear is stimulated by more than one tone, sometimes other tones are heard which are not part of the stimulus. These "subjective" tones became known as "combination tones."

Despite many challenges, Helmholtz' theory rapidly became widely accepted because of its simplicity and comprehensiveness (see reviews by Rayleigh,[13] Schouten,[14] Plomp[15]). Many investigators gathered new psychophysical and physiological evidence relevant to his hypotheses; we shall mention only a few. Wegel and Lane[16] studied the masking effects of tones upon other tones. They found that the ear does not behave as a set of resonators tuned as sharply as Helmhotz thought, but that there was general support for his first hypothesis, in that masking effects are larger when tones are closer together in frequency. Von Békésy, starting in 1941, directly observed the mechanical properties and vibration patterns of the basilar membrane in response to simple tones and found that they can be described as traveling waves, having a rather broad envelope maximum whose location depends on the stimulus frequency (for a comprehensive review, see von Békésy[17]). Later electrophysiological studies of the auditory nerve showed that at that point the system is definitely tonotopically organized (Galambos and Davis,[18, 19] Tasaki,[20] Kiang[21]).

Von Békésy[22, 23] also tried to find support for Helmholtz' second principle in a somewhat more generalized form. He tried to reconcile the rather broad tuning characteristics of the cochlea with the ear's ability to detect very small differences in frequency by postulating a neural sharpening model which he tried to simulate with the tactile sense of the skin. He observed that a very broad pattern of stimulation can result in a strongly localized sensation, which he considers as support for a possible neural sharpening process in the place-to-pitch transformation.

Helmholtz' third hypothesis concerning auditory nonlinearity has also obtained support from later psychophysical and physiological work, with important modifications: (a) combination tones can occur at all sound intensities; (b) quadratic summation and difference tones are less important than Helmholtz thought; and (c) the cochlea, rather than the middle ear, is the primary generator of combination tones. Helmholtz' viewpoint that combination tones behave like tones added to the stimulus has been fully supported, however, (Zwicker,[24] Plomp,[15] Goldstein,[25-28] Goldstein and Kiang[29]).

Fletcher[30] also tried to reconcile Seebeck's experimental observations quantitatively with Helmholtz' theory. He showed that for a periodic sound a tone corresponding to the fundamental can be heard even when all lower harmonics (up to about the 7th are filtered out of the stimulus. He used Helmholtz' combination-tone hypothesis to explain this effect: Whenever energy at the fundamental frequency is missing in a periodic sound, the ear will reintroduce it by nonlinear distortion.

From the beginning, Helmholtz' concepts immediately evoked valid criticism. Stumpf[31] found it difficult to reconcile the dimensions of the cochlea with the dimensions necessary to obtain resonance for the lower audible frequencies. König[32] produced amplitude-modulated sounds by using a tuning fork and a siren disk. He reported a tone sensation corresponding to the interruption rate, and later Hermann[33] found that this tonal sensation can vary up to 20% when the frequency of the tone is not an integer multiple of the interruption rate. For this last phenomenon Hermann had no explanation, but for the "harmonic" case he argued that the ear must perceive the periodicity in a sound without the necessity of energy at that frequency. This was obviously in disagreement with Ohm's law. Physiological experiments with nerve-muscle preparations showed that stimulating the nerve with a periodic pulse could cause a vibration and an audible tone in the muscle, indicating that nerve fibers could carry timing information (Helmholtz,[34] Bernstein[35]). Thompson[36] confirmed a claim made earlier by Dove[37] that when two tones of slightly different low frequencies are presented dichotically, one to each ear, a clear beat is heard, which shows that under certain circumstances periodicity information is preserved in the auditory nerve. On the basis of these findings, Wundt[38] argued that it is the rate of nerve impulses, synchronous to the stimulus tone, that gives rise to a tonal sensation, and not the place of an active nerve fiber or group of fibers. This idea was later formulated and developed by Wever,[39] and became known as the "volley theory." Theories of hearing built exclusively on this periodicity principle, and denying Helmholtz' resonance theory altogether, were proposed by Rutherford[40] and by Meyer.[41-45]

Schouten[46,47] re-established Seebeck's original observations and demonstrated for the first time that these cannot be reconciled with Helmholtz' combination-tone and place-pitch theories as Fletcher had tried to do. Using an optical siren, he generated a periodic pulse train in which the fundamental frequency was cancelled. He observed that even at relatively low intensities this signal sounded similar to the original pulse train, and not an octave higher. Upon adding a small test tone near the frequency of the missing fundamental he found that no beats were produced, and showed that a distortion tone could not be responsible for his and Seebeck's observations. The growing evidence of a tonotopic organization in the cochlea and the auditory nerve kept him, however, from rejecting Helmholtz' resonance principle; his theory was an attempt to reconcile the resonance and periodicity concepts. Schouten's hypotheses are (a) The ear performs a frequency analysis with limited resolving power, in accord with Helmholtz' first hypothesis; (b) At each place, the periodicity of the resulting waveform is preserved "in the periodicity of the excitation of the receptors," which is "transmitted by the nerve fibres" and mapped into a pitch sensation; and (c) The place of excitation determines the "tone quality," sharpness or timbre.

This model predicts that for a complex periodic signal the lower harmonics will be perceived as separate tones, limited by the resolving power of the cochlear filters. The higher harmonics, which are not resolved, result in a waveform whose envelope periodicity is the same as that of the fundamental frequency of the stimulus. This periodicity is preserved in the connected nerve fibers and causes a tonal sensation which Schouten[48] calls "the residue." His theory explains Seebeck's observation that higher harmonics, sharing a common period, somehow add to the sensation of the fundamental.

De Boer[49] confirmed an earlier observation by Small[50] that for a fixed number of harmonic components the tonality decreases with increasing mean frequency. He also studied the effect of inharmonic tone complexes, which Hermann and Schouten had reported on much earlier but had not explained. He postulated two mechanisms, one operating in the temporal domain, the second in the spectral domain:

1. For complex tones with closely spaced frequencies the tonal sensation is derived from the fine-structure temporal periodicity generated by (partially) unresolved stimulus tones.

2. For complex tones with widely spaced frequencies the sensation is derived from an approximate common divisor to the stimulus frequencies.

Schouten, Ritsma, and Cardozo[51] emphasized the fine-structure periodicity mechanism and developed an empirical formula relating the pitch of a three-tone AM complex to its constitutent frequencies. Like de Boer, they found consistent discrepancies between their experimental results and the predictions from the fine-structure peak-to-peak distances in the AM stimulus waveform. They defined the simple theoretical prediction as the "first effect," and the measured discrepancy as the "second effect." The spectral formulation also leads to a second effect. Fishler[52] proposed a model (equivalent to de Boer's time formulation) consisting in linear superposition of the stimulus frequencies, followed by a peak-to-peak measurement in the time domain. Such a model does qualitatively produce a second effect, but not large enough to explain the data of de Boer[49] and of Schouten and others.[51] Walliser[53] proposed an operation called "division and approximation," wherein the musical equivalent of an inharmonic complex is given by that subharmonic of the lowest stimulus frequency which is closest to the difference frequency. The predictions of this model are the same as de Boer's and Fishler's models when all weight is put on the lowest frequency in the complex, and hence Walliser's model does not produce enough of a second effect to explain the experimental data satisfactorily. The empirical second effect suggested that for higher harmonic numbers the effective spectral components had to be below the stimulus components. Schroeder[54] had proposed that a certain amount of phase modulation be added to the amplitude modulation, thereby effectively adding spectral components that are not present in the stimulus. Goldstein and Kiang[29] found physiological evidence that combination tones are part of the effective stimulus; therefore, they suggested that it is reasonable to account for the extreme second effect with an effective signal that contains significant spectral energy at lower frequencies than exist in the actual stimulus. Ritsma,[55] who studied the pitch of three-tone complexes, and Smoorenburg,[56] who dealt with two-tone complexes, concluded from their data that the effective harmonic number is never larger than approximately 9, and that combination tones provide an explanation for the extent of the second effect.

The finding concerning the upper bound of the effective harmonic number was consistent with earlier studies by Ritsma.[57-60] He found that the "tonal residue" for a fundamental of ~200 Hz does not extend beyond the 10th or 15th harmonic, depending on the experimental paradigm. Following Schouten's model, he then searched for a dominant place or frequency region where a periodicity measurement is made, and found that this is given by the 3rd to 5th harmonics.

It is interesting to notice that these last results, although obtained in studies that embraced Schouten's residue hypothesis, implied a contradiction of that hypothesis. According to Schouten, periodicity pitch sensation should be bounded by a minimum harmonic number, determined by the ear's resolving power. Later studies seem to indicate, however, that the real bound is an upper bound which is quantitatively not much different from the ear's spectral resolving power, according to behavioral studies on resolution of stimulus tones (Plomp[15]) and aural combination tones (Goldstein[26]).

Other mechanisms than those mentioned thus far have been proposed. Some investigators have claimed that the residue effect is very weak or based upon mental faculties that are not auditory-specific. Thurlow[61] observed that subjects often respond with great hesitancy when presented with a complex sound comprising upper partials only. They seemed to use vocal humming as an intermediate response. This made him conclude that the listener does not "hear" the residue as a percept, but matches percepts of the partials of his voice and those of the stimulus sound, and extrapolates the "best fitting" fundamental. A related argument can be advanced that normally complex tones preserve the fundamental melody information in the frequency ratios of upper partials; tracking a particular partial of fixed harmonic number would recover the melody described by the missing fundamental. Many of the results of tone-matching experiments by various investigators can be explained as matching of spectral components if we consider that subjects have a good estimate of frequency ratios other than just the unison (Ward,[62a] Houtsma[62b]).

When we consider our original observations of the ear's ability to track a melody in a series of complex periodic sounds in the light of this brief historical review, we conclude that there is little disagreement about the fact that what the ear tracks is the fundamental frequency. The question of how the ear extracts and encodes the information of the fundamental, however, is still very much unsettled. The principal contrasting viewpoints concerning this last issue are, we believe:

1. Energy at the fundamental frequency is a necessary condition; if it is not present in the stimulus, the ear produces it by nonlinear distortion (place-pitch theory).
2. Fundamental periodicity at some place in the cochlea is a necessary condition (periodicity-pitch theory).
3. No percept corresponding to the fundamental exists when the stimulus has no energy at that frequency. Rather, a fundamental is associated with the sound via some form of higher mental processing of percepts produced by the partials.

The purpose of this study is twofold: (i) to investigate the phenomenon of fundamental tracking in an experimental paradigm that closely simulates musical behavior; and (ii) to produce new qualitative and quantitative experimental evidence in the light of which the unsettled concepts mentioned above can be evaluated.

1.1 SUMMARY

In Section II we shall discuss some qualitative results of experiments on melody perception and recognition, using stimuli comprising 2 successive harmonics. Such stimuli are rather easy to generate and much is known about the physiological responses that they evoke in the auditory nerve (Goldstein and Kiang,[29] Sachs and Kiang,[63] Sachs[64]). Section III will cover quantitative experiments on musical-interval identification, using two-tone stimuli, both monotic and dichotic, and will show the need for a common central mechanism underlying the fundamental tracking phenomenon. These experiments explore in detail the salient properties revealed in Section II; they represent the heart of this report. In Section IV we shall examine the existence region of this phenomenon and its dependence on the number of partials in the stimulus, while phase effects will be discussed in Section V. Section VI contains experiments using inharmonic two-tone stimuli, both monotic and dichotic, and provides additional evidence that the phenomenon under study is the same as that described in psychoacoustical literature as "the residue." In Section VII we attempt to apply a simple decision model to reduce the data of Section III to a single sensitivity parameter, and an empirical formula is derived to describe its dependence on some relevant stimulus parameters. Conclusions and some suggestions for future research are mentioned in Section VIII.

1.2 EQUIPMENT

The equipment used for most of the experiments was three programmable oscillators (one GR 1161-A frequency synthesizer and two KrohnHite 4031R oscillators), two Grason & Stadler electronic switches, a dual audio amplifier, and a set of TDH 39 headphones. Harmonic and intermodulation distortion measured at the headphones was better than 50 dB below stimulus level. In some experiments, when more than three tones were needed simultaneously, a dual multiplier (CBL 47) was used. Subjects were seated in an IAC Model 1200 sound-insulated chamber. Switches and oscillators were controlled

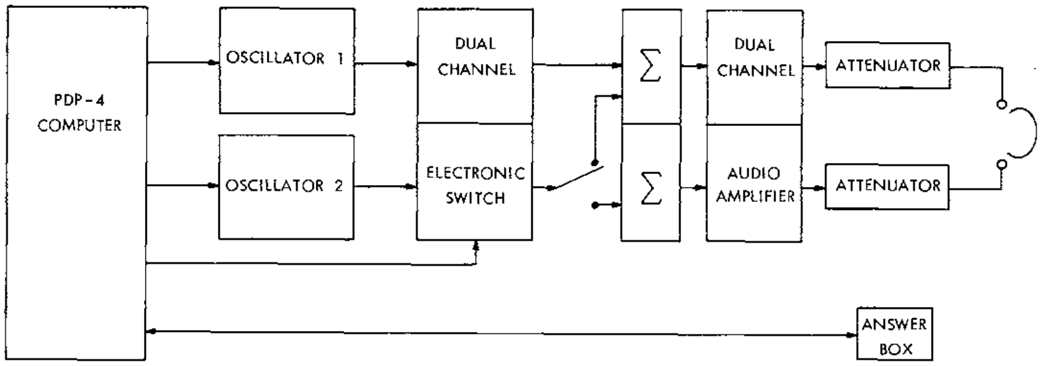


Fig. 1. Equipment diagram for all two-tone experiments.

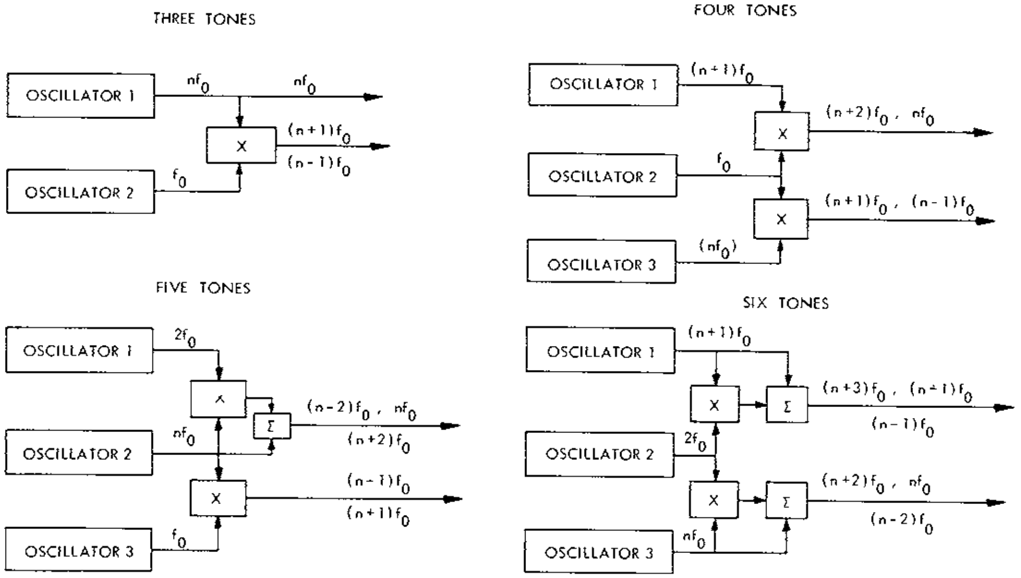


Fig. 2. Equipment diagram for all experiments using more than two stimulus tones.

by a DEC PDP-4 computer, which generated all random events, performed stimulus computations, stored responses and controlled feedback. Equipment diagrams are shown in Figs. 1 and 2.

1.3 SUBJECTS IN EXPERIMENTS

Ten subjects, seven men and three women, participated in the qualitative experiments described in Section II. All of them had at least some degree of musical training and were familiar with musical notation and dictation. For most subjects training proved unnecessary; a few were given some practice, starting with complexes of 6 harmonics which were then gradually reduced to the two-tone stimuli used in the main experiments.

Three subjects from these ten were selected to participate in the rest of the experiments on the basis of availability, interest, and general performance. All of them were men, 18 to 31 years old, who had quite extensive musical training and experience (majors in organ, viola, and singing). No special training sessions were given, and test sessions were limited to a maximum 2 hours daily.

II. PRELIMINARY EXPERIMENTS ON MELODY RECOGNITION USING TWO-TONE STIMULI

2.1 EXPERIMENT 1

We have discussed general musical behavior, in which the ear can track a melody in a sequence of periodic sounds, rather independently of the number of harmonics or the relative energy at particular harmonics. Ritsma[65] and his associates showed in their demonstration phonograph record that placing a formant filter at different center frequencies does not affect the perceived melody, but only the tone quality or timbre. It should be noticed though that in experiments of this kind, as well as in music played on conventional instruments, the spectral formant is either stationary or moves parallel with the fundamental frequency. It is not obvious, however, what the ear would perceive when the fundamental and the spectral formant move independently of each other. Cross and Lane[66] have found evidence that prior training determines the cue to which the ear attends. Therefore we performed an experiment in which subjects

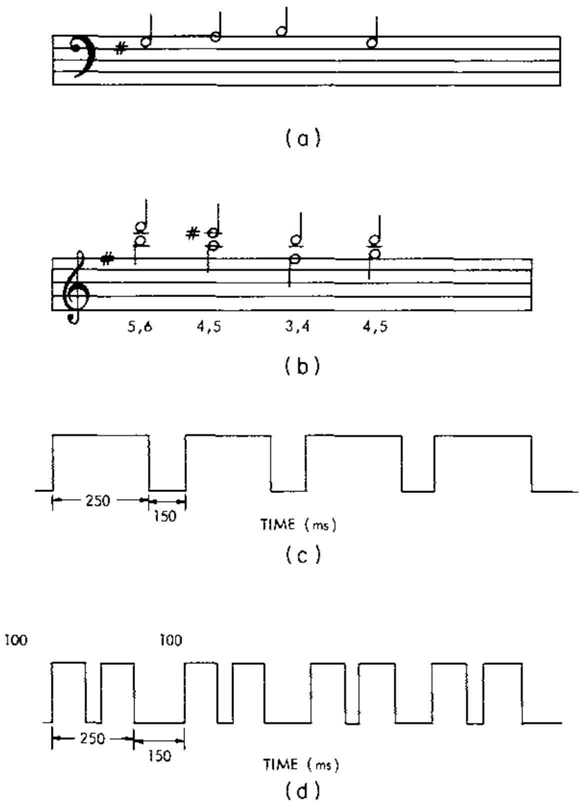


Fig. 3. Experimental paradigm for four-note melody experiments. (a) Typical fundamental melody. (b) Sample of actual stimulus frequencies; labels indicate particular harmonic numbers. (c and d) Time structure of stimuli for melody experiments.

were given a sequence of 4 periodic sounds with fundamental frequencies between 200 Hz and 400 Hz, describing a four-note melody as illustrated in Fig. 3. The sound was just two successive harmonics, the numbers being varied systematically over a limited range from note to note. They could be the 3rd and 4th, 4th and 5th, or 5th and 6th harmonics of a fundamental corresponding to a particular note. A sample is shown in Fig. 3b. Stimuli were presented monotically at a sound pressure level (SPL) of 50 dB. The time envelope of the sound sequence is shown in Fig. 3c. The subjects, all of whom were familiar with musical notation, were asked to write what melody they heard. Some reported a melody that was consistent with the missing fundamental, as in Fig. 3a; others reported a melody corresponding to the partials as in Fig. 3b. One subject was able to report both melodies described by the upper and lower partials.

Helmholtz[9] mentioned that complex periodic sounds can be perceived "synthetically" (that is, the complex is perceived as one sound having one pitch) or "analytically" (that is, partials are heard individually, each one having its own pitch). Cross and Lane[66] showed that these two modes of behavior can be controlled by previous training. They had two groups of subjects scale 25 periodic stimuli, which were combinations of 5 fundamental frequencies and 5 formant filters, in 5 categories after training with only 5 stimuli. Depending on the choice of the five training stimuli, the subjects scaled consistently according to the fundamental frequency or the location of spectral energy. We therefore tried in a second experiment to get all subjects to operate in the "fundamental mode" by eliminating all relevant information in individual partials.

2.2 EXPERIMENT 2

This experiment was similar to the first, except that there were 4 melodies similar to the one illustrated in Fig. 3a. The subject had control over which of the four melodies was to be presented by pushing one of 4 buttons on an answer box. The harmonic numbers for each note were not chosen systematically, as in the first experiment, but at random, over the same range as in Experiment 1. Subjects were instructed to listen to each melody several times, and report a "consistent feature," if it could be heard, on a relative note scale. Then, in the second part of the experiment, 50 trials were presented comprising the same four melodies in random order, again with random harmonic numbers, and the subjects were asked to identify each melody by pushing one of the four buttons. After each presentation they were given 3 seconds in which to respond, after which feedback was provided by means of an electronic number display indicating the correct melody number. This experiment can be characterized as an identification experiment of 4 classes determined by the sequence of fundamentals, each class having 81 elements (3 possible choices of harmonic pairs per note).

After little training, the results were that 9 of the 10 subjects tested characterized each class (corresponding to each button on the box) by the missing fundamentals in the first part, and scored perfectly or nearly perfectly in the second part. These results demonstrate a musical phenomenon of fundamental tracking, that, under certain conditions, a musical message comprising a series of fundamental periods can be retrieved aurally from a sequence of periodic sounds, irrespective of their harmonic content.

2.3 EXPERIMENT 3

As we have pointed out, there are many ways in which the fundamental frequency could be derived from a pair of successive harmonics. This experiment and Experiment 4 were designed to examine whether some of several possible operations which the auditory system could perform can be eliminated. The experimental paradigm was the same as in Experiment 2, except that in addition to choosing a random pair of harmonics for each note, both harmonics were shifted up or down in frequency by a random but equal amount, limited to one-fourth of the fundamental frequency, thereby keeping the difference frequency equal to the fundamental but making the complex tone inharmonic. Stimuli were presented at a sound pressure level of 50 dB and the subjects were given the same instructions as in Experiment 2. Note that in this experiment each class has virtually an infinite number of elements because of the random continuous frequency shift. It turned out that even after repeated attempts none of the subjects could hear any consistent melody in the sequences of 4 sounds; hence, no subject could give a note description for any of the four classes. The identification part, like that in Experiment 2, was therefore omitted. These results show that information of the original fundamental was not perceived, despite the fact that it was preserved by the transformed sounds in the form of the difference frequency.

2.4 EXPERIMENT 4

In this experiment we examined whether a necessary condition for fundamental tracking is that both harmonics be present at the same time. The experimental paradigm was the same as that of Experiment 2, except that instead of both harmonics being presented simultaneously, they were presented in time sequence, as illustrated in Fig. 3d. The subjects were instructed as in Experiment 2, and most of them knew how the stimuli were organized. Despite considerable practice, none of them was able to determine the fundamental frequencies or musical intervals underlying each class of stimuli, even after having heard a large number of samples from each particular class. Most of the subjects agreed that more practice would not help and that the task could not be done. Therefore we conclude that the simultaneous presence of more than one harmonic is indeed a necessary condition for fundamental tracking.

2.5 EXPERIMENT 5

In Experiments 1-4 we have seen, among other things, that 2 successive harmonics with numbers between 3 and 6, simultaneously presented to one ear, create a sensation that is musically equivalent to a tone of the missing fundamental frequency. From the viewpoint of the residue theory this seems paradoxical, since harmonics of low order can easily be resolved behaviorally, given the appropriate experimental paradigm (Plomp[15]). Also, on the basis of known cochlear electrophysiology (Kiang,[21] Sachs and Kiang,[63] Goldstein and Kiang,[29] Goldstein[27]), insignificant cochlear interaction would be expected for such low-order harmonics. This further suggests that there is some kind of central processor that receives its inputs for each resolved tone through separate channels. This separate-channel hypothesis was tested by the extreme channel separation of using separate ears. We thought that, although negative results would not necessarily invalidate the hypothesis, positive results would undeniably prove it to be correct. The experimental paradigm was that of Experiment 2, except that both harmonics were presented dichotically, one to each ear. The performance of 6 of 10 subjects was the same as in the monaural version. They reported melodies corresponding to the missing fundamentals and scored nearly perfectly in the identification part. Three subjects had some difficulties in both parts, which probably could have been improved by more practice; this was not given, however. One subject, who also had serious difficulties in the monaural experiment, was unable to perform the task. The exact percentage of the subjects who could perform this task, the scores of those subjects who found this experiment more difficult than others, and how their performance might improve with more training are issues that are not considered to be very important. The performance of most of the subjects tested does show that the missing fundamental can be tracked reliably from pairs of successive harmonics presented through separate ears.

2.6 DISCUSSION

Through a series of qualitative and objective experiments we have established the phenomenon of fundamental tracking and have investigated some conditions under which it does or does not exist. By qualitative we mean that the only question we were trying to answer was whether a subject could or could not perform a particular task. Matters such as the exact number of right and wrong answers, the number of occasional mistakes in tracking the fundamental, the amount of training required for each subject, and the possible contribution of previous musical experience were not considered, although some of these issues may be worthwhile to study. By objective we mean that the subject was asked a question and was required to answer, the answer being either right or wrong. The subject's performance was then taken as a criterion for the presence or absence of the phenomenon under study, while the experimenter tried to control systematically those features of the stimulus which might provide the subject with information about the correct answer.

When we compare the experimental results obtained from Experiments 1-5 with the different theories described in Section I, we see that none of these theories is sufficient to explain all of the results. Helmholtz' theory that a melody can be perceived only if there is energy at the respective fundamentals is clearly contradicted by the results of Experiment 2, in that subjects were able to track the fundamentals while no energy was present at these frequencies. The possibility that such energy is reintroduced in the ear by nonlinear distortion is rendered highly improbable by the fact that these experiments were carried out at a stimulus intensity of 50 dB SPL, at which the difference tone is below the threshold of perception (Goldstein[26]), and by the results of Experiment 3. If difference frequency per se were a relevant clue for the subject, his performance should not depend on random shifts in stimulus frequencies, as long as their difference is kept unchanged. The possibility of difference tone distortion is finally eliminated categorically by the results of Experiment 5, in which the notes were heard with only one tone in each ear.

Periodicity theories, in their extreme form (Meyer) or combined with peripheral frequency analysis (Schouten), do indeed predict the results of Experiments 1, 2, and 4, and also with some modifications of Experiment 3. The results of Experiment 5 cannot possibly be explained by any form or modification of the periodicity theory. With dichotically separated tones there is no place in the peripheral auditory system where envelope periodicities, corresponding to the missing fundamental, could possibly arise.

Some of the theories in which it is assumed that there is no direct perceptional counterpart of the missing fundamental but that the fundamental is indirectly derived from perceptions of upper partials (viewpoint 3 in Section I), are not clearly supported by our experiments, but neither can such theories be categorically rejected. We thought that if the fundamental were derived from independent sensations of individual partials, the paradigm of Experiment 4 would make the task of fundamental tracking easier for the subject, rather than more difficult. The results show, however, that under these conditions the task is not only more difficult but completely impossible, which proves the necessity of the simultaneous presence of at least two partials. Since the duration of the stimulus (Fig. 3c) and the response times were relatively short in all of our experiments, it appears that there would not be sufficient time to perform elaborate computations such as those involving humming (Thurlow[61]). We know of no way of proving that a direct percept of the fundamental exists; introspective reports from our subjects that "the melodies were really there" are the only information that we have about this matter. All of our experiments were performed from the viewpoint of measuring musical behavior, and in this report questions of the existence or absence of a percept are considered moot.

The use of dichotic experiments to differentiate between peripheral and central effects is not new. Experiments on binaural beats were mentioned in Section I. Dove[37] used dichotically presented partials to prove that Tartini tones (combination tones arising from monaural distortion) are not subjective but objective. He failed to hear a difference tone when two successive partials were presented to different ears through rubber tubes. Similar experiments were performed later by Thompson.[36, 67, 68] Although both investigators noticed the absence of a distortion tone in the dichotic case, they did not notice, or at least did not report, any musical pitch sensation corresponding to the missing fundamental.

Our dichotic experiments prove directly that neither fundamental energy nor fundamental periods in the cochlear output are necessary conditions for fundamental tracking. This is a new finding; it shows that, at least for the dichotic conditions, fundamental tracking behavior must be mediated by a central, neural mechanism operating on signals derived from separated partials. The location of this mechanism can be anywhere at or beyond the level of the superior olivary complex.

III. IDENTIFICATION OF MUSICAL INTERVALS USING TWO-TONE STIMULI

3.1 INTRODUCTION

The results of the experiments described in Section II, specifically the dichotic experiment, demonstrate the necessity for a central neural mechanism with bilateral inputs. This in no way disproves the existence of other mechanisms which may be employed when stimulus tones are presented monotically. In order to determine whether both monotic and dichotic fundamental tracking can be accounted for by one central mechanism, or whether additional mechanisms are required, we carried out a series of quantitative experiments. Comparison of the results from monotic and dichotic paradigms could indicate whether or not a common mechanism is sufficient.

3.2 EXPERIMENTS – GENERAL REMARK

Five experiments that have many common features were performed. First, these features will be described. Then each experiment will be described and the results will be discussed.

The subjects' task was to identify on each given trial which one of 8 known two-note melodies (or musical intervals) was presented. These simple melodies had identical envelope time structure and all began with the same note. The notes were tuned to the natural scale, with frequency ratios of 16/15, 9/8, 6/5, and 5/4 for the minor and major second, and minor and major third, respectively. The stimuli representing the notes comprised two successive harmonics, the number of the lower harmonic being chosen randomly for each note over a range of 3 successive integers. The middle of the range of lower harmonic numbers, $\bar{n}$, and the fundamental frequency of the first note, f_0, were chosen as independent parameters in measuring identification performance, expressed in percentage correct responses. The basic experimental paradigm is illustrated in Fig. 4. The reason why this particular set of musical intervals was chosen is not that we assume an inherent significance in the intervals, but merely that, at least in Western culture, they provide a convenient language which musically trained people can understand. Whether in other cultures a different choice of intervals would lead to similar results is beyond the scope of this work. The three subjects who participated in this series of experiments were first tested for their ability to identify musical intervals played with square-wave stimuli at fundamental frequencies equal to the notes in Fig. 4a; control runs of 50 trials were carried out. Subjects were given a key similar to Fig. 4a which ascribed a number to each interval, and were instructed to push a corresponding number on their answer box within 4 seconds after each stimulus presentation, after which feedback was provided. All three subjects scored perfectly.

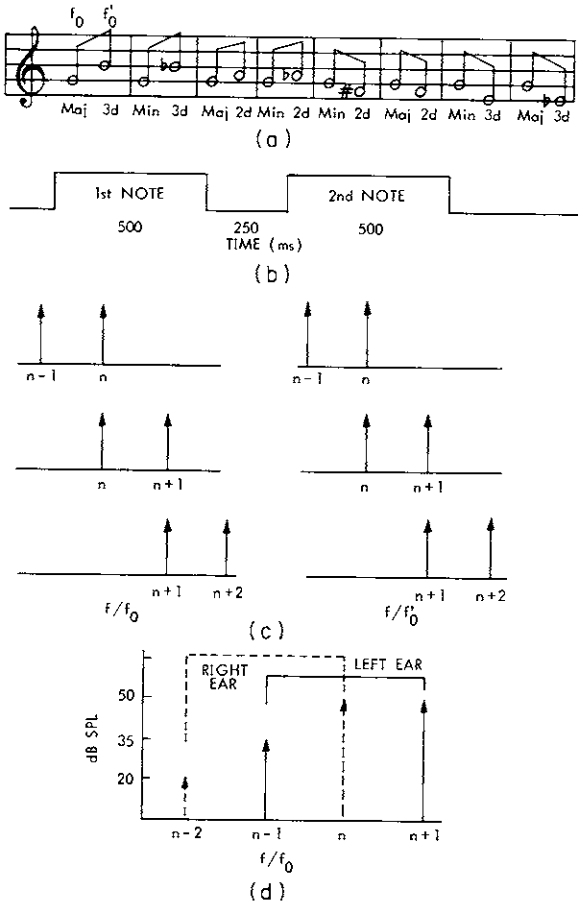


Fig. 4. Experimental paradigm for musical-interval identification experiments.
(a) Musical intervals to be identified.
(b) Time envelope of the total stimulus.
(c) The three possible two-tone stimuli for each of the two notes; for each note a random choice was made among the three possible stimuli.
(d) Stimulus configuration for one note of a dichotically presented melody, with the addition of simulated combination tones.

In Experiments 6-10, the subjects were tested individually; A. H. and S. W. were always given feedback, but N. H. was not because he found it distracting. Typical runs were 50 trials for Subject A. H., and 25 trials for Subjects S. W. and N. H. Fundamental frequencies were taken in steps of 100 Hz or sometimes 200 Hz, as indicated in the figures; for each fundamental frequency several runs were taken with increasing average harmonic number $\bar{n}$. As many runs were taken as were necessary to make performance drop from perfect identification to essentially chance. (One out of eight correct is chance response.) Then a psychometric function was fitted by eye for each fundamental frequency, $\bar{n}$ being the independent variable, and from these functions equal performance contours were plotted. These psychometric functions are plotted in Appendix I.

3.3 EXPERIMENT 6

In Experiment 6 stimuli were presented at a 20-dB sensation level, the intensities being adjusted for each value of $\bar{n}$ to roughly match each subject's audiogram. The audiograms are shown in Fig. 5. Contours of equal performance from this experiment are shown in Fig. 6 for each subject. Additional data for an extended fundamental frequency range were obtained from only one subject (A. H.) and are shown in Fig. 11a. On Subject A. H. a control experiment was performed with the partials of each note presented in time sequence, similarly to Experiment 4. Values for f_0 and $\bar{n}$ were chosen which had yielded perfect performance for monotic simultaneous presentation. Despite great effort, Subject A. H. was not able to perform better than chance, from which we conclude that randomization of harmonic number forces the subject to use both partials, and that these partials must be presented simultaneously.

3.4 EXPERIMENT 7

This experiment was made in the same way as Experiment 6, except that the two simulus tones were presented dichotically, one to each ear. Results are shown in Fig. 7 and Fig. 11b.

A quick comparison of Figs. 6 and 7 shows that each subject's performance is essentially identical in the monotic and the dichotic tests, which suggests that indeed a central mechanism integrates and processes information from both cochleas, and that the inputs to this mechanism are similar under monotic and dichotic stimulation.

3.5 EXPERIMENTS 8 AND 9

To further investigate the similarities in performance for monotic and dichotic conditions, Experiments 8 and 9 were undertaken at a higher stimulus intensity level, 50 dB SPL for Subjects A. H. and S. W., and 40 dB SPL for Subject N. H., monotic and dichotic, respectively. Equal performance contours for the monotic and the dichotic experiments are shown in Fig. 8 and Fig. 9, respectively.

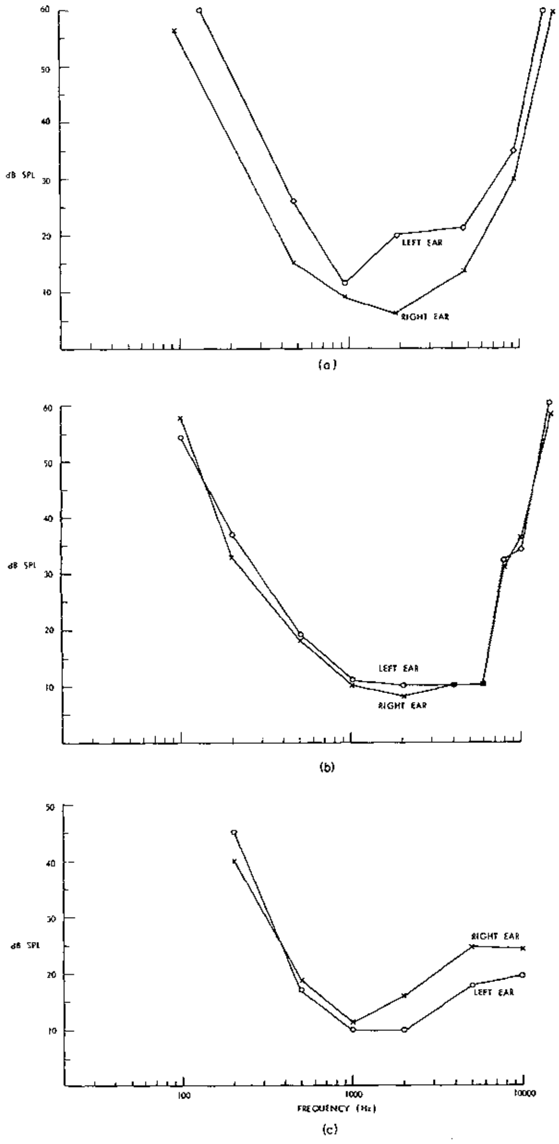


Fig. 5.

Audiograms. (a) Subject N. H. (b) Subject A.H. (c) Subject S.W.

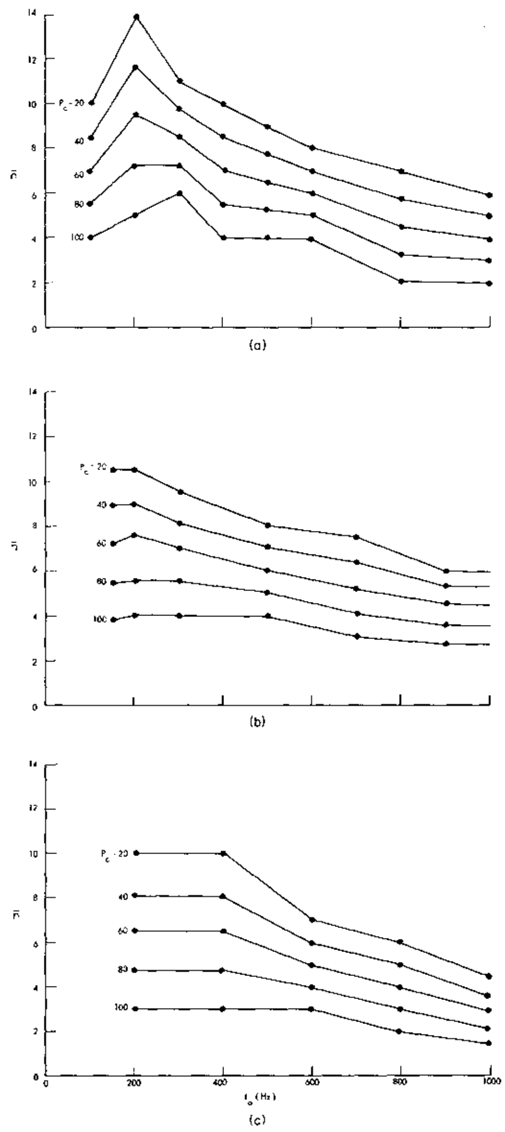


Fig. 6.

Performance contours, monotic, 20 dB SL. (a) Subject N. H. (b) Subject A. H. (c) Subject S. W.

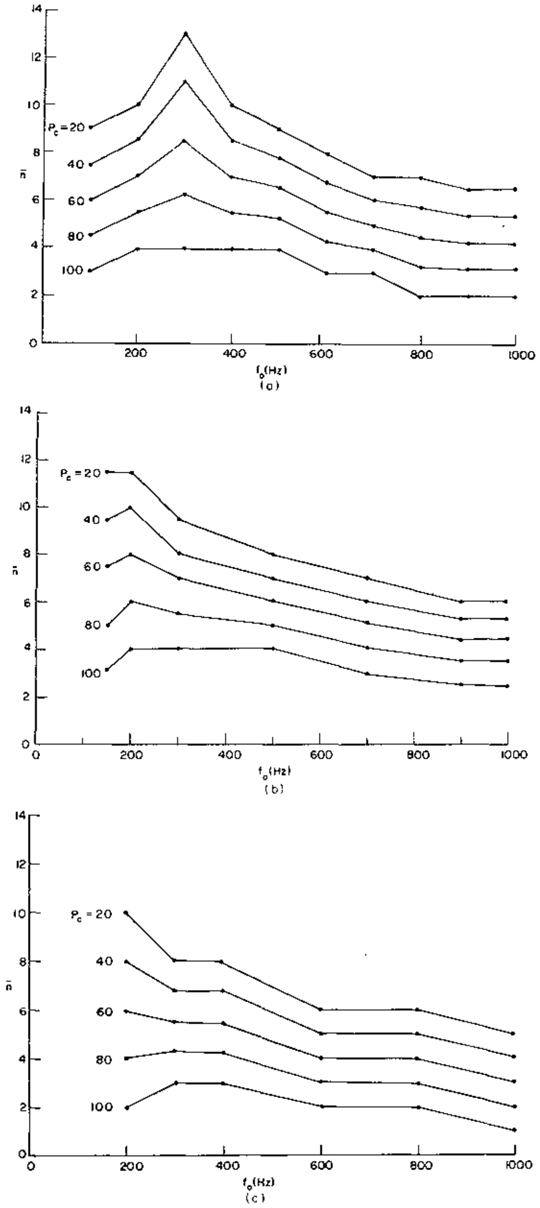


Fig. 7.

Performance contours, dichotic, 20 dB SL. (a) Subject N.H. (b) Subject A.H. (c) Subject S.W.

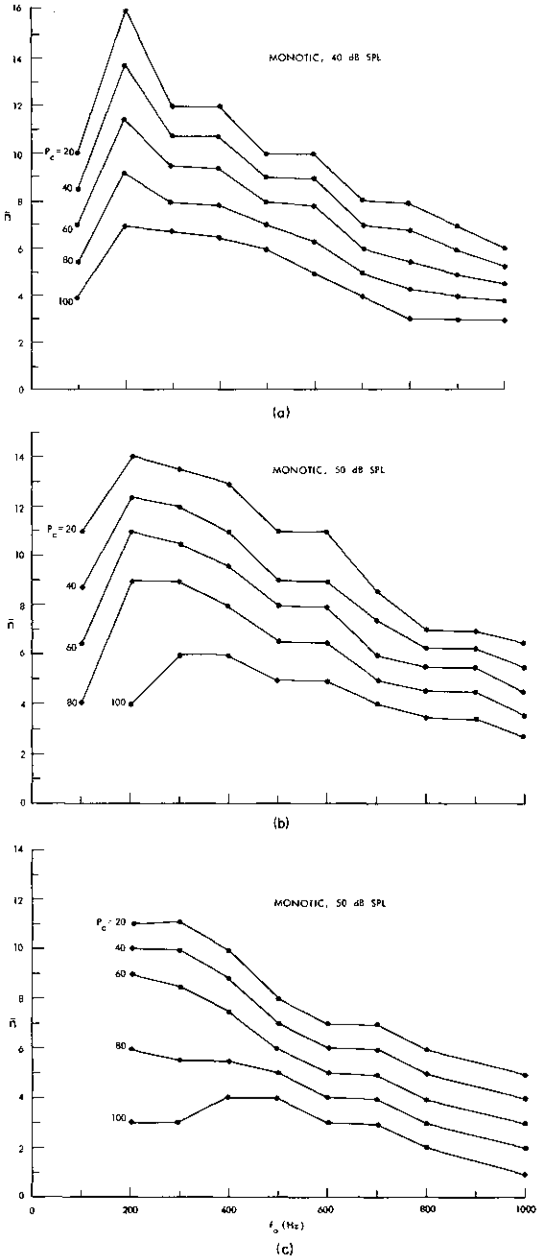


Fig. 8.

Performance contours, monotic, 40 and 50 dB SPL. (a) Subject N.H. (b) Subject A.H. (c) Subject S.W.

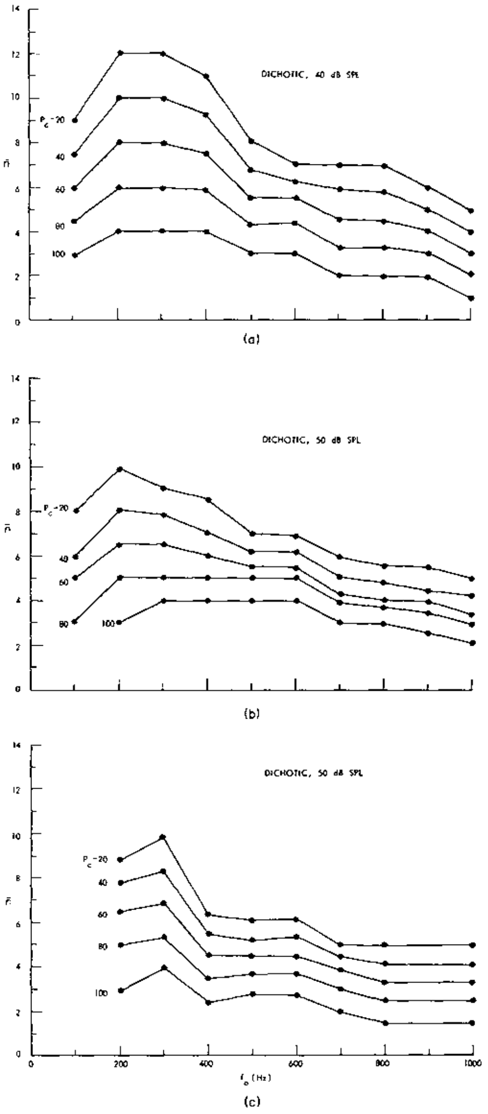


Fig. 9.

Performance contours, dichotic, 40 and 50 dB SPL. (a) Subject N. H. (b) Subject A. H. (c) Subject S. W.

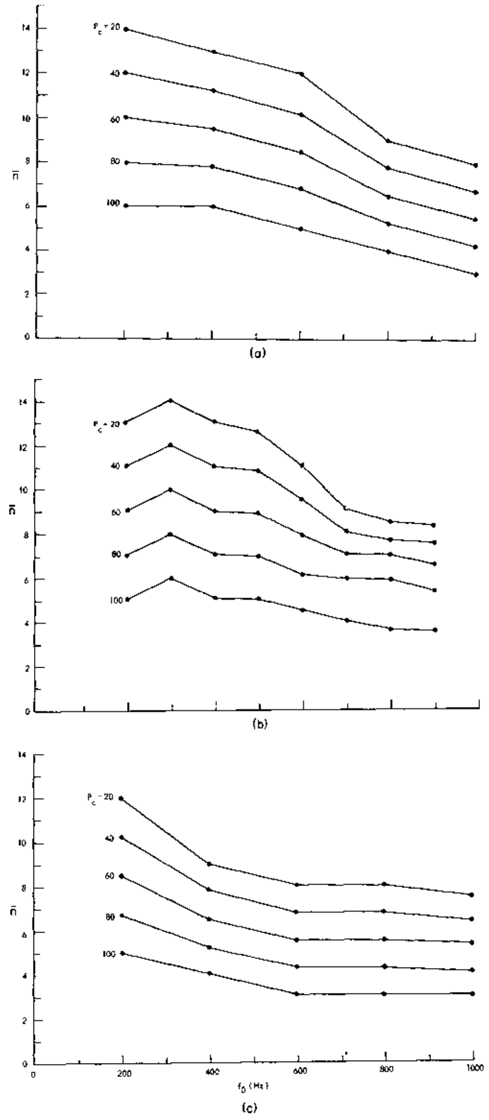


Fig. 10.

Performance contours, dichotic, 40 and 50 dB SPL, simulated combination tones added. Subjects and primary levels as in Fig. 9.

Comparing results from all four experiments, we see that each subject's performance is essentially the same under all four stimulus conditions, except that for higher intensity monotic stimulation (Fig. 8) performance contours are shifted upward by approximately 2 or 3 harmonic numbers. Such an upward shift might be expected because of the presence of aural combination tones of type $f_1-k(f_2-f_1)$ generated in the peripheral ear for a monotic stimulus comprising the frequencies f_1 and f_2 (Goldstein,[26] Goldstein and Kiang[29]). These combination tones provide the ear with 2 or 3 harmonics below those contained in the stimulus, which are probably very useful because all results thus far indicate that performance improves with decreasing harmonic number. These combination tones could make the effective average harmonic number approximately 2 or 3 lower than the actual value of $\bar{n}$ in Fig. 8. In the dichotic experiments combination tones are not present, and in the monotic experiment at 20 dB SL they would be near or below threshold.

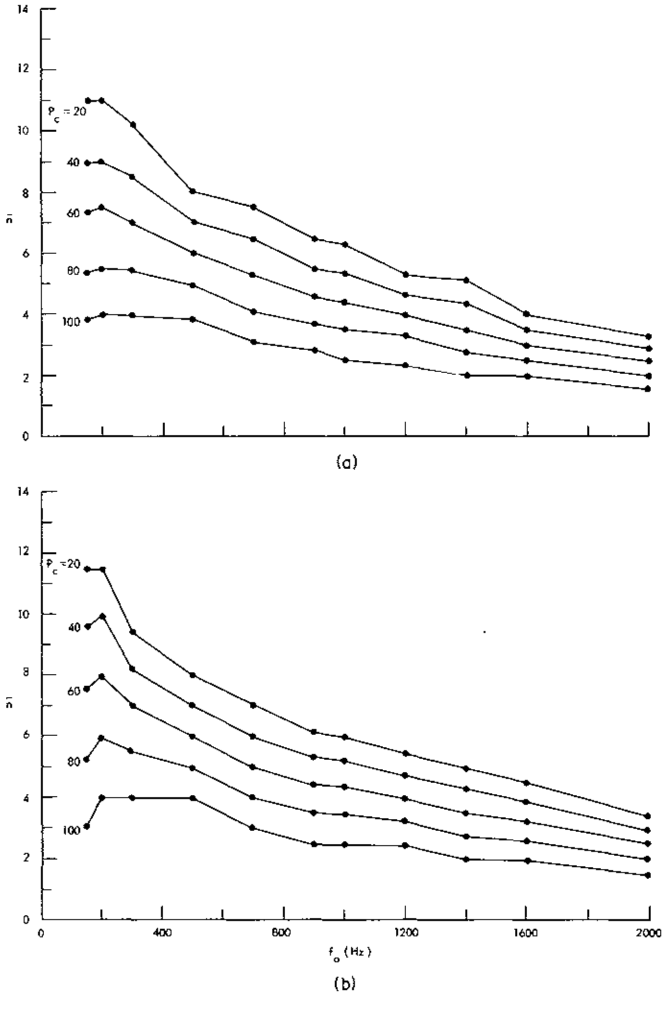


Fig. 11. Performance contours for Subject A. H. (a) monotic, 20 dB SL. (b) dichotic, 20 dB SL.

3.6 EXPERIMENT 10

To test this combination-tone hypothesis more directly, Experiment 10 was carried out using dichotic stimuli at 50 dB SPL (40 dB for Subject N.H.) with the addition of 2 tones that approximately simulate the aural combination tones that the ear generates under monotic conditions. The stimulus paradigm is shown in Fig. 4d, and the experimental results in Fig. 10. The same upward shift of performance contours as in Experiment 8 can be readily noticed, and the similarity of Figs. 8 and 11 furnishes strong evidence that the performance differences that do occur between monotic and dichotic stimulus conditions can be attributed to combination tones generated in the peripheral ear.

3.7 DISCUSSION

The results of this series of experiments show that all theories that have been developed to explain how we can track a melody in a sequence of complex sounds are inadequate. A place-pitch theory based on Helmholtz's principles would predict for the monotic paradigms a performance of approximately 41% correct for all harmonic numbers up to the point where the ear's frequency resolution limit is reached. If we assume that a musical interval described by resolvable partials in the two-note sequence can always be recognized perfectly, then at least one-third of all trials must be identified correctly because the harmonic numbers of the second note will be the same as those of the first, and both partials will form the same interval as the missing fundamental. Then, on the average, one-eighth of the remaining trials will be answered correctly by chance, making the average correct score approximately 41%. It is clear from our data that this prediction was not borne out. We might wonder why performance for the dichotic paradigms was ever less than 41%, since under these conditions there should not be a frequency resolution bound as in the monotic case. The answer may be that the subjects did not or could not, because of physiological constraints, switch their strategy from listening for all intervals "in the same key" to attempting to hear intervals described by partials in transposed keys.

A periodicity model like that developed by Schouten would predict a behavior just opposite to that of our experimental results. The lower harmonics, which are resolved in the cochlea, would not provide the right kind of information about the fundamental because they are randomly chosen. Only the higher harmonics, which cannot be separately resolved, would reflect the fundamental in their envelope periodicity, irrespective of their exact harmonic number. Hence the model would predict that correct identification should increase with $\bar{n}$, while the data clearly show that the opposite is true over the whole fundamental frequency range.

Our results in general are consistent with Ritsma's empirical finding that harmonics 3 to 5 are the most dominant in providing a musical equivalent to the fundamental.

Our data do not show, however, that performance deteriorates for harmonic numbers lower than 4; performance for stimulus parameters below the 100% correct contour is always perfect. It might be that Ritsma's experiments were more sensitive than ours in this respect, and that we missed some small effect.

Smoorenburg's conclusion that the effective harmonic numbers for fundamentals of 200 Hz have an upper bound of approximately 9 is also supported and extended to other fundamentals by our results. For $\bar{n}$ greater than 9, performance is essentially chance for any fundamental frequency, unless combination tones provide effective harmonics below this upper bound, as with the 50-dB SPL monotic stimuli. This upper bound also coincides rather closely with the limit to the ear's frequency resolving power (Plomp[15]). It is interesting that the same boundary applies when the stimulus tones are presented dichotically. Obviously, limited frequency resolution in the peripheral ear cannot be responsible for restricting perception of the missing fundamental to tones of harmonic number below approximately 10; the cause must be more central.

The question put at the beginning of this section about whether or not one common neural mechanism is sufficient to account for all fundamental tracking behavior has been answered unambiguously by our experimental findings. The qualitative and quantitative similarity of the data for monotic and dichotic conditions and for different intensities eliminates the need for more than one fundamental tracking mechanism. The contour shift for medium-intensity monotic stimulation can be well accounted for by aural combination tones added by the peripheral ear.

In Section II, it was shown that neither energy at the fundamental frequency nor fundamental periodicity in the cochlear outputs is a necessary condition for fundamental tracking. From a converse point of view we can now state that a sufficient condition is given by energy at the fundamental, which can easily be shown experimentally, but probably not by fundamental periods in the cochlear output. The experimental data from Experiments 6-10 suggest that such fundamental periods are probably irrelevant for the following reasons: (a) Monotic stimuli, which can provide cochlear fundamental periods, give no better performance than dichotic stimuli. (b) With monotic stimuli the possibilities for fundamental periods in the cochlear output are enhanced as the harmonic number $\bar{n}$ is increased. Yet interval identification performance deteriorates with increasing harmonic number. (c) All monotic stimuli for which identification performance was better than chance were either behaviorally resolvable tones or stimuli that generated these tones as combination tones (Plomp,[15] Goldstein[26]).

The close correlation between the limits on fundamental tracking and behavioral frequency resolution and the similarity of monotic and dichotic performance suggest that for complex-tone stimuli the fundamental tracking mechanism operates on those stimulus tones or combination tones that can be resolved by the cochlea. This conclusion is a radical departure from the theory of the "residue," which is defined as "the joint perception of those higher Fourier components which the ear fails to resolve" (Schouten, Ritsma, Cardozo[51]).

IV. FUNDAMENTAL TRACKING AND THE NUMBER OF STIMULUS PARTIALS

4.1 INTRODUCTION

The experiments described in Sections II and III employed stimuli comprising only two successive harmonics. When we perceive a melody from a musical instrument, we are usually presented with a much larger number of harmonics. Most studies investigating the auditory system's ability to extract information of the fundamental from a complex sound employed signals containing more than two partials. Seebeck[3] and Schouten[46] used pulse trains, which contain a very large number of harmonics. The same can be said of Thurlow and Small[69] and of Flanagan and Guttman.[70, 71] De Boer[49] studied pitch effects of harmonic and inharmonic tone complexes comprising 5 and 7 partials; he thought that 5 was about the lowest number from which a stable and distinct "residue" could be obtained. Schouten and others[51] showed that 3 successive upper partials, contained in an AM complex, are sufficient to evoke a fundamental sensation, and that from a behavioral point of view such signals are equivalent to those used by de Boer. Recently, Smoorenburg[56] showed that the same is true for two-tone signals. Nevertheless, there seems to be general agreement and some direct experimental evidence (Walliser[53]) that the sensation of the missing fundamental becomes stronger with an increasing number of upper partials. In order to investigate how behavoir depends quantitatively on the number of harmonics present in a stimulus, and whether experimental results from stimuli comprising various numbers of partials all reflect basically the same phenomenon, Experiments 11 and 12 were carried out.

4.2 EXPERIMENT 11

Using the interval identification paradigm described in Section III, we studied recognition behavior for stimuli comprising 3 successive harmonics. Stimuli were generated by means of two oscillators and a modulator so that the partials had equal intensity and were in AM phase. The total stimulus intensity was kept at a sensation level of 20 dB. For monotic presentation the stimulus partials were added and presented to one ear, and for dichotic presentation the carrier was presented to one ear, the side tones to the other ear. Runs of 25 trials were taken for two subjects, S.M. and R.C., both women, who had had wide musical experience (singing). Figures 12 and 13 show equal performance contours for monotic and dichotic stimuli, respectively. The variable $\bar{n}$ represents the center of the range of the lowest partial; as in the experiments in Section III, partials could vary randomly for each note over a range of 3 harmonic numbers.

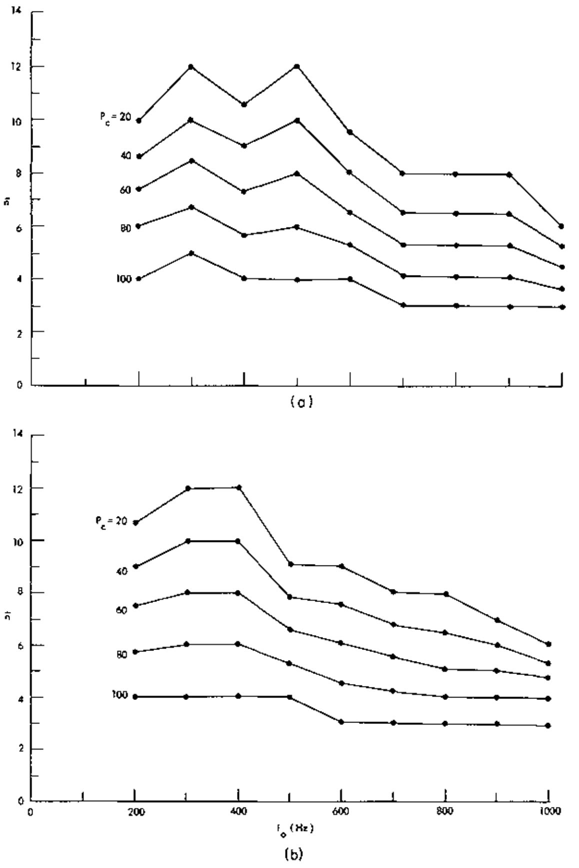


Fig. 12.

Performance contours, 3 harmonics, monotic, 20 dB SL. (a) Subject R. C.; (b) Subject S. M.

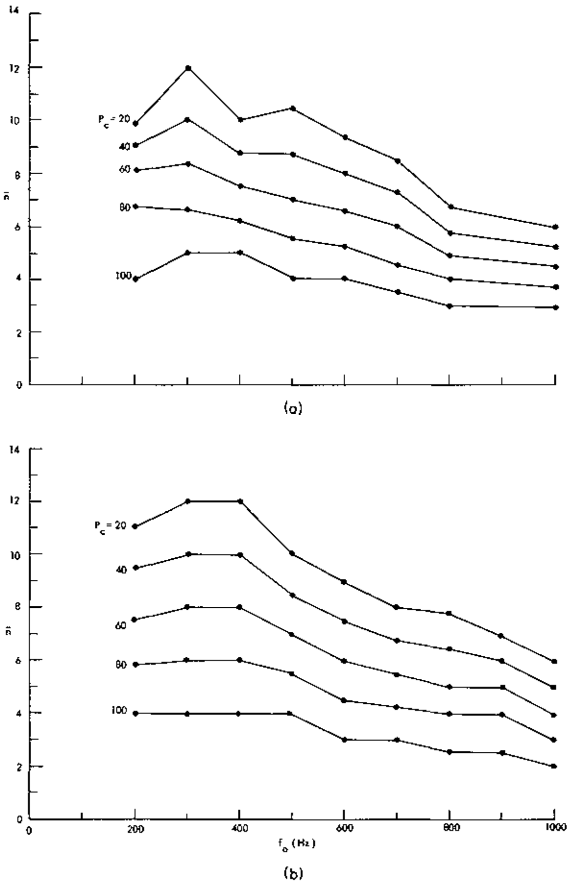


Fig. 13.

Performance contours, 3 harmonics, dichotic, 20 dB SL. (a) Subject R. C.; (b) Subject S. M.

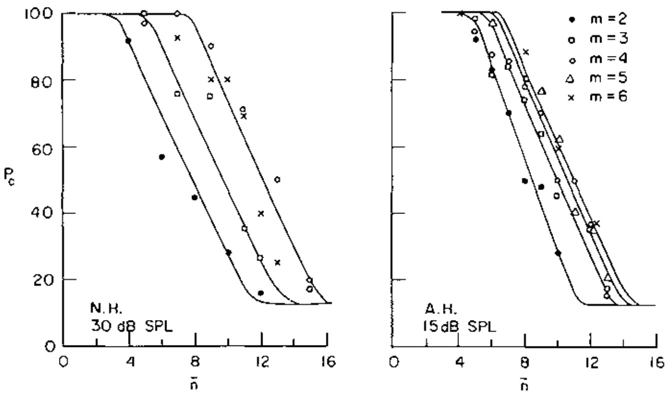


Fig. 14. Psychometric functions for 2 subjects in an eight-alternative musical-interval identification experiment, using stimuli comprising m harmonics.

4.3 EXPERIMENT 12

Using the same interval identification paradigm as in Section III, performance was studied for just one fundamental frequency (300 Hz) as a function of $\bar{n}$, the center of the range of the lowest partial. The number of partials in the stimulus, m, was an experimental parameter and the range over which the lowest of the m successive harmonics was randomly chosen for each note was extended from 3 to 5. Stimuli were presented at sensation levels of 15 dB for Subject A. H. and 30 dB for Subject N. H. In order to minimize possible combination tones, the stimulus partials were divided dichotically into even and odd harmonics, so that each ear was never stimulated by two successive harmonics. Each run included 50 trials for Subject A. H. and 25 for Subject N. H. A family of psychometric functions for each subject is shown in Fig. 14.

4.4 DISCUSSION

The results of Experiment 11 show great similarity with identification data for stimuli comprising only two partials. Behavior for monotic and dichotic stimulus conditions is essentially identical, and the general tendencies are the same as for two-tone stimuli. Further comparisions should not be made, because of the different subjects that were employed in the three-tone experiments. Experiment 12 enables us to compare performance of two subjects for stimuli containing various numbers of harmonics. It shows that fundamental tracking performance does improve with an increasing number of partials contained in the stimulus, which is consistent with the notion that we have mentioned that the more harmonics the stimulus contains, the stronger is the sensation of the fundamental. But the data show also that this is true only for a limited range of m; when the stimulus contains 4 harmonics or more, the addition of higher harmonics will not add anything behaviorally useful to the fundamental sensation.

The largest value of m in our experiments was 6. Ritsma's finding[58] that for a bandpass-filtered pulse train with a pulse rate of 280 pps the maximum harmonic number of the lowest audible harmonics for which subjects could just hear a "tonal residue" was 11 or 12 is consistent with our findings; the width of his bandpass filter was such that the stimulus contained a large number of harmonics.

The results of both experiments, which involve more than two partials, are also consistent with the notion that a central processor operates on those stimulus partials that are peripherally resolved. Specifically, we point to the similarity in the performance of both subjects in Experiment 11 under monotic and dichotic stimulus conditions, and at the similar bounds on the lowest stimulus partials in Experiment 12 in which dichotic stimuli were used and in Ritsma's monotic results.

The fact that in Experiment 12 one subject (N. H.) still performed significantly better than chance with four- or six-tone stimuli and a value of 13 for $\bar{n}$ does not necessarily contradict our condition of peripheral frequency resolution for the following reasons:

1. The range of random choice of the lowest harmonic number was extended to 5 in this experiment; this means that when $\bar{n}$ equals 13, 40% of the trials will have a lowest harmonic number of 11 or 12.

2. Since for this particular subject stimuli were presented at a sensation level of 30 dB, and the partials were distributed between the two ears in such a way that it is quite likely that combination tones similar to type $f_1 - k(f_2-f_1)$ for two-tone stimuli played an important role. For instance, when the total stimulus comprises harmonics 13 through 18, one ear receiving harmonics 13, 15, and 17 may generate harmonics 11 and 9 as audible combination tones, and the other ear may similarly produce harmonics 12 and 10. The fact that aural combination tones can indeed extend the range of fundamental tracking beyond the point where stimulus partials are known to be behaviorally resolved, has been demonstrated in Section III.

In summary, the main conclusions from the experimental results described in this section are as follows.

1. All fundamental tracking behavior in harmonic tone complexes comprising two or more harmonic frequency components reflects the same basic phenomenon.

2. Within the range of peripherally resolvable harmonics, a larger number of harmonics in a tone complex will result in better fundamental tracking performance up to a certain point. When the number of such harmonics exceeds approximately 4, the improvement effect saturates.

3. For all harmonic complex stimuli, regardless of the number of stimulus partials, a necessary condition for fundamental tracking is that some of the lower partials of the <u>effective</u> stimulus, including possible aural combination tones, can be peripherally resolved, independently of whether partials are presented monotically or dichotically. From the dichotic two-tone experiments described in Section III it is clear, however, that peripheral resolution is not in itself a <u>sufficient</u> condition for fundamental tracking because each ear receives only one tone and yet performance clearly decays with increasing harmonic number.

V. EFFECTS OF RELATIVE PHASE IN TONE COMPLEXES COMPRISING MORE THAN TWO HARMONICS

5.1 INTRODUCTION

We have shown the necessity for a central neural "pitch processor" receiving separate inputs derived from peripherally resolved stimulus components. When more than two harmonic components are employed, the system shows the same general behavior as for two-tone stimuli. When a stimulus comprising more than two harmonics is resolved into its partials in the cochlea, and the cochlear filtering and neural transformation processes retain the information of the phase relations between stimulus components, then it is quite possible that fundamental tracking behavior is dependent on the relative phases of the stimulus partials.

Mathes and Miller[72] reported a clearly perceptible change in sensation when the carrier in a monotically presented AM complex was changed from a 0° phase relation (AM phase) to a 90° phase (QFM phase). The former caused the complex to sound quite harsh, accompanied by a low pitch corresponding to the modulation frequency, while a much smoother sound resulted from the latter phase relation. They noted that the harsh quality of the sound was always observed when the stimulus waveform showed a maximum of amplitude modulation, and they suggested a relation between their observations and Schouten's "residue." Licklider[73] used this argument to explain why Hoogland[74] was unable to observe a "residue" in a complex sound generated by a number of carefully tuned oscillators. Since Hoogland did not control the phase relations between harmonics, on which the "residue" presumably depends, he did not observe it. Licklider demonstrated that a low pitch, corresponding to the missing fundamental is easily heard when the relative phases are adjusted in such a way that the stimulus waveform has a maximally impulsive pattern. This finding was re-established later by other investigators, among them de Boer[49] and Walliser.[53]

Ritsma and Engel[75] studied the pitch of a harmonic three-tone complex with the carrier in quasi FM phase by having subjects adjust the missing fundamental of a harmonic three-tone AM signal until the pitches matched. They found with this procedure that the pitch of the QFM complex showed significant and systematic deviations from its fundamental frequency. Thus the low pitch that can be heard in a harmonic three-tone complex can not only be made more or less pronounced by changing the carrier phase, but actually can be changed.

The observation that the perception of a low pitch depends heavily on the relative phase relationships between stimulus components is contradicted by Ritsma's studies[59, 60] on the "dominant frequencies" in complex sounds. If indeed the 3rd, 4th and 5th harmonics are most prominent in creating a fundamental pitch sensation, then they are beyond the limit of phase sensitivity of the ear to monaurally presented tone complexes because they are more than 20% apart (de Boer,[49] Goldstein[25, 26]). In other words, the most prominent partials for fundamental pitch perception are just those to which the ear is phase-insensitive!

Our findings that fundamental tracking behavior is basically the same for monotically and dichotically presented tone complexes enables us to examine the question of the influence of relative phase on fundamental perception directly by means of a simple experiment.

5.2 EXPERIMENT 13

A two-interval, two-alternative forced-choice discrimination experiment was carried out between two three-tone complexes, one in AM phase and the other in QFM phase, for both monotic and dichotic stimulus conditions. In the monotic case all three tones were presented to one ear, while in the dichotic case the carrier and side tones were presented to different ears. The carrier frequency, f_c, was kept at 2000 Hz and at twice the amplitude of the side tones, thereby producing a 100% modulation depth for the AM phase. The modulation frequency, f_m, was varied in steps from 25 Hz to 500 Hz. Stimulus intensity was kept at 40 dB SPL. Only one subject (A. H.) participated in this experiment. In each trial he was presented with a 500-ms soundburst of the AM phase signal, shortly followed by a 500-ms burst of the QFM signal, or in reverse order. The signals differed only in the carrier phase. He had to make a binary choice in each trial about the correct order in which the signals were presented. Feedback about the correctness of the response was provided after each trial. An equipment diagram is shown in Fig. 15. For several values of the modulation frequency f_m, runs of 25 trials were taken for both monotic and dichotic stimulus conditions. The resulting psychometric functions are shown in Fig. 16.

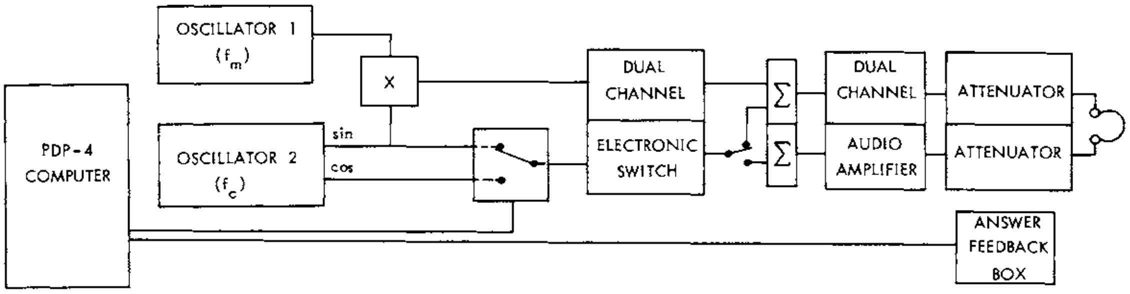


Fig. 15. Equipment diagram for two-alternative forced-choice phase discrimination experiments. The bipolar switch is so programmed that for each trial the probability of either state is 50%.

It is quite evident that for monotic stimuli there is a strong phase effect. For partials spacings less than ~12% the AM and QFM signals are perfectly discriminable; when the spacing is more than 20%, they apparently sound identical, as indicated by the chance-level performance. The psychometric function is not a simple monotonic one. The subject reported that a clue reversal took place in the following sense: With increasing

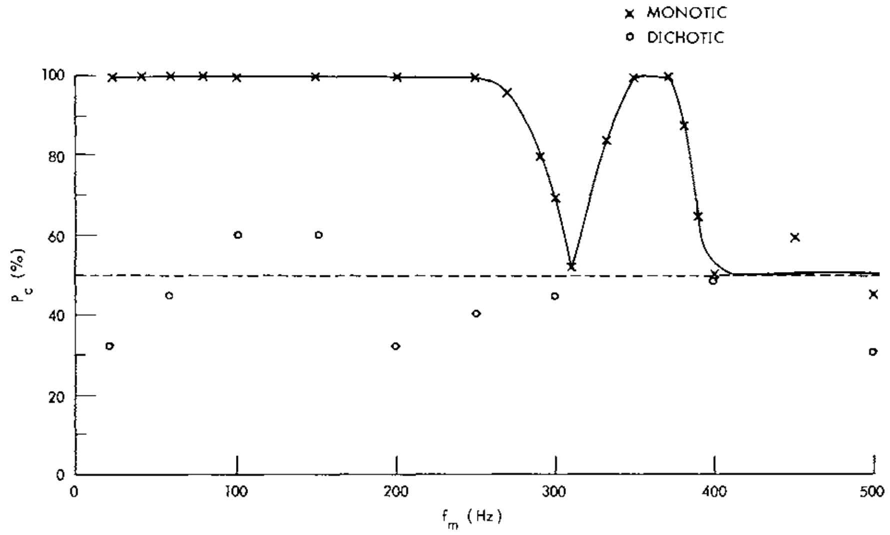


Fig. 16. Psychometric functions for a two-interval, two-alternative forced-choice phase discrimination experiment, using three-tone complexes. Stimuli were presented monotically and dichotically at 40 dB SPL. Carrier frequency f_c, 2000 Hz. Subject A. H.

f_m, the AM and QFM complexes were initially distinguishable by their harsh and smooth sounds, respectively, then they began to sound more and more similar, after which they became again easily distinguishable by smooth and harsh qualities, respectively, and finally they became indistinguishable for good. This subjective report is consistent with the subject's performance reflected in a "bi-modal" psychometric function.

From the experimental results it is equally evident that for dichotic presentation, the carrier in one ear and both side tones in the other ear, there is no noticeable phase effect. No modulation frequency f_m was found for which performance was significantly better than chance. Several other carrier frequencies and stimulus intensities were tried, but none of them yielded positive results.

5.3 EXPERIMENT 14

Another, more qualitative, experiment was carried out. It is well known that if we present an AM suppressed carrier signal monotically and add a slightly mistuned carrier to the signal, a clear beat will be heard, provided the spacing between the signal components is smaller than ~20% (Kelvin,[76] de Boer[49]), since a mistuned carrier can be interpreted as if it were precisely tuned but with a phase angle that linearly increases with time. If the signal components are spaced in such a way that AM and QFM phase relations can easily be distinguished aurally, then some kind of beat-like effect should be expected when the signal changes back and forth between these states in a continuous way. It has been observed experimentally that a clear beat is heard also at exactly those stimulus frequencies and intensities that yielded a "dip" in the psychometric function for the forced-choice experiment. This indicates that the dip should not be interpreted to mean that at these stimulus frequencies the system is insensitive to phase, but only that it is insensitive to the difference of the two particular phase relations. We made an attempt to create a similar beatlike effect with the slightly mistuned carrier in one ear and the side tones in the other ear. A large number of carrier frequencies, modulation frequencies, and intensities were tried with 3 subjects, and no combination was found for which beats or beatlike effects were reported.

5.4 DISCUSSION

These experiments give a clear answer to the question posed at the beginning of this section. Fundamental tracking behavior with two- and three-tone stimuli at low intensities is essentially identical for the monotic and dichotic paradigms. Thus we conclude that the central pitch processor treats monotic and dichotic signals equivalently. We have shown that dichotic three-tone stimuli do not give noticeable phase effects. Therefore we conclude that the central pitch processor is insensitive to the relative phase relations between its separate inputs. This means that the monaural phase effects demonstrated in one of our experiments and reported elsewhere (Goldstein[25, 26]), and their relation to the sensation of the "residue," must have a more peripheral cause. A possible explanation could be given by the cochlea's nonlinear properties. Goldstein[27] reported that for a monaural stimulus comprising 3 equally distant frequency components f_1, f_2, and f_3, the intensities of the aural combination tones of type $f_1-k(f_2-f_1)$ were very greatly dependent on the relative intensities and phase relations of the stimulus tones. The exact relations have not been mapped out quantitatively, so that there is no way, at present, to show, for instance, that the dependence of combination tones on relative phase of stimulus frequencies does explain the results of our monotic phase experiment. It is possible that aural combination tones and related mechanisms may be sufficient to explain many of the established monaural phase effects.

The conclusion that the central pitch processor is phase-insensitive means that Licklider's account of Hoogland's failure to hear a "residue" is only a partial explanation. Indeed, by adjusting the phase of stimulus components, we might maximize some intensities or an average intensity of aural combination tones, which then might add to the sensation of the missing fundamental, but this can only be a secondary effect. Hoogland should have been able to hear a "residue" since we have shown that subjects can track a missing fundamental from dichotically presented harmonics where phase is irrelevant. The primary reason for Hoogland's failure is that he chose harmonic numbers in the vicinity of approximately 30; the results of Section III show that with such a value of $\bar{n}$ the fundamental cannot be tracked.

Ritsma and Engel's report that phase changes in the carrier of a three-tone complex caused large pitch changes may appear to be basically inconsistent with our conclusion that the pitch processor is insensitive to the phase relations among its inputs. Their results, however, could also be explained by peripherally generated combination tones. The harmonic numbers which they used were larger than 9, which is just about at the limit of behavioral frequency resolution. Combination tones, which have been shown to depend on the phase of the carrier, can conceivably change the pitch impression by drastically changing the spectral composition of the *effective* stimulus. They report also that when harmonic numbers around 5 or 6 were chosen, the pitch of the QFM complex was always judged equal to the missing fundamental, as for an AM phase carrier, which is completely consistent with our findings and conclusions.

The conclusion drawn from Experiments 13 and 14 is that the extraction of fundamental information from a tone complex is not directly dependent on the relative phase of the tones. The central neural processor is insensitive to the phase relations of its separate inputs. The established phase effects for monaurally presented tone complexes must therefore be caused by some mechanism peripheral to this central processor. Cochlear nonlinearities that generate combination tones play a major role in producing these phase effects.

VI. AURAL TRACKING OF INHARMONIC TWO-TONE COMPLEXES

6.1 INTRODUCTION

Thus far we have focused our study of musical behavior on periodic stimuli, namely complex tones whose partial frequencies are exact multiples of one and the same number which was called the "missing fundamental." This restriction enabled us to refer to note identification in a sequence of such sounds as "fundamental tracking." When one or more stimulus partials are shifted in frequency from a harmonic situation, the complex tone becomes inharmonic and we can no longer speak about a missing fundamental, since the component frequencies are no longer integer multiples of some number. One might question whether or not such an inharmonic sound has a specific musical pitch in the operational sense that it can be associated with a note on the musical scale. We shall make an attempt to answer this question. The results will provide a logical link with earlier work on musical pitch.

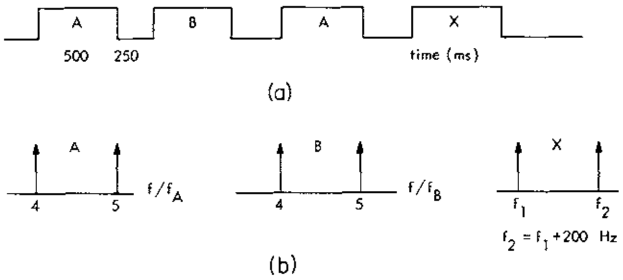
Experiment 3 dealt with a specific case of inharmonic sounds. In that experiment no melody was perceived when for each note of the melody score the pair of successive harmonics was shifted in frequency by a random amount. This finding does not answer the question of whether the inharmonic tone complexes do not have a definable musical pitch or whether they do have a pitch that depends lawfully on the random shift, so that the original melody would have been transformed into an unrecognizable melody.

Several experiments have been reported that offer a definite answer to this question. Hermann[33] and Schouten[14, 48] noticed that shifting the frequencies of harmonic partials from a harmonic situation causes a change in pitch. De Boer,[49] Schouten, Ritsma, and Cardozo,[51] Walliser,[53] and Smoorenburg[56] made systematic studies of the pitch of inharmonic complex tones by having subjects adjust a parameter of a comparison sound, either the missing fundamental of a harmonic tone complex or the frequency of a simple tone (Walliser[53]), until the inharmonic test sound and the comparison sound seemed to be tuned to the same note. They found systematic relationships between the stimulus frequencies of the inharmonic complex and the adjustable parameter of the comparison sound, presumably reflecting its note value or musical pitch. These relationships have been discussed as the first and second effects of inharmonic frequency shifts. They show that making a complex stimulus inharmonic by introducing a frequency shift does not eliminate its musical pitch but merely changes it. They also demonstrate that this pitch is not determined by frequency difference per se.

In our study of fundamental tracking we have used very different experimental procedures from the matching procedures used by these authors. Despite our departure from earlier procedures, we believe that our investigations involved the same phenomenon as theirs. This belief is further supported by the results of the following experiments.

6.2 EXPERIMENT 15

A matching experiment was performed in which a two-note melody A-B was matched to another melody A-X. A and B were harmonic two-tone complexes with fixed harmonic numbers of 4 and 5. X had two simple tones with a spectral difference of 200 Hz. The fundamental of B, f_B, was set equal to one of the seven frequencies at 5-Hz steps in the range 185-215 Hz, and f_A, the fundamental of A, was always a full tone below that of B (that is, a frequency ratio of 8/9). Stimuli were presented at sound pressure levels of 40 and 60 dB as indicated in the graphs; the paradigm is illustrated in Fig. 17. The subject had continuous control over the frequency f_1, the lower of the two partials of the complex tone X; the higher frequency component, f_2, was always 200 Hz larger than f_1. Only one subject (A. H.) participated in this experiment; he aurally matched the intervals of A-X and A-B by adjusting the frequencies of X. Three sets of data were obtained, each point representing 2 settings. Two of the sets were for monotic stimulus presentation and the third for dichotic conditions in which the stimulus partials were presented to different ears. The data are shown in Fig. 18. All three sets of data can be described in a first approximation by straight lines drawn through the harmonic frequencies f_1 = 200n Hz and f_B = 200 Hz, with a slope of 1/n.



f_A (Hz)	f_B (Hz)
191	215
187	210
182	205
178	200
173	195
169	190
164	185

(c)

Fig. 17. Experimental paradigm used in Experiment 15.
(a) Time envelope of the stimulus.
(b) Spectral composition of stimulus components.
(c) Fundamental frequencies for tone complexes A and B.

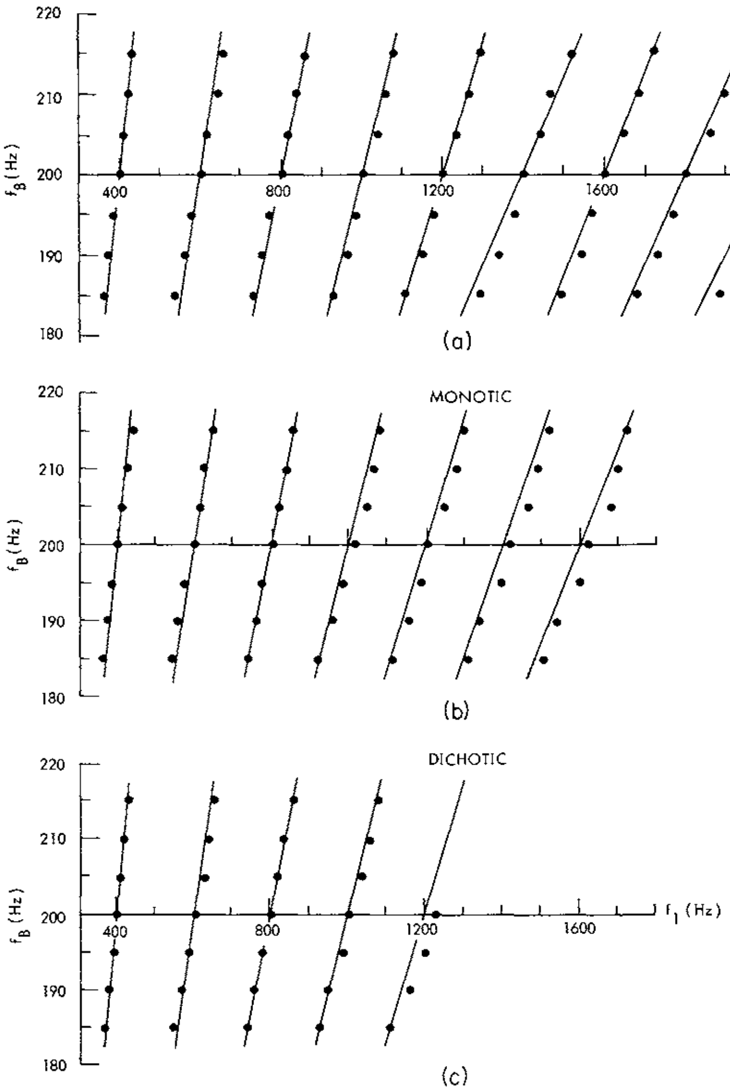


Fig. 18. (a) Results of the A B A X musical-interval matching experiment using the paradigm shown in Fig. 17. Tones were presented monotically at 60 dB SPL. Data are roughly described by straight lines having a slope of 1/n. Subject A. H.
(b and c) Stimulus tones presented monotically and dichotically at 40 dB SPL. Otherwise the same as in (a). Subject A. H.

We judged that more than two replications would be required for a more precise description of the data than a first approximation; the first approximation, however, provides sufficient detail for present purposes.

6.3 EXPERIMENT 16

In Experiment 15 the harmonic numbers of the reference stimulus (B) were kept the same while the inharmonic complex X was derived from harmonic situations with various values of the harmonic number n. It was seen that the relation between inharmonic frequency shift away from a harmonic position and the fundamental of a harmonic comparison complex tone is roughly linear with a slope dependent on n, the harmonic number of the complex tone from which the inharmonic stimulus is derived. The slope, however, could also depend on the harmonic number of the complex reference tone. Therefore an experiment was performed similar to Experiment 15, but always starting the complex X from the harmonic situation $f_1 = 600$ Hz, $f_2 = 800$ Hz. The harmonic number n of the complexes A and B were varied systematically from 2 to 8 (both being the same), and for each value a set of matching data was obtained using the same values for f_A and f_B as those shown in Fig. 17. Only monotic data were taken for one subject (A. H.); data points, representing 2 settings each, are given in Fig. 19. A straight line

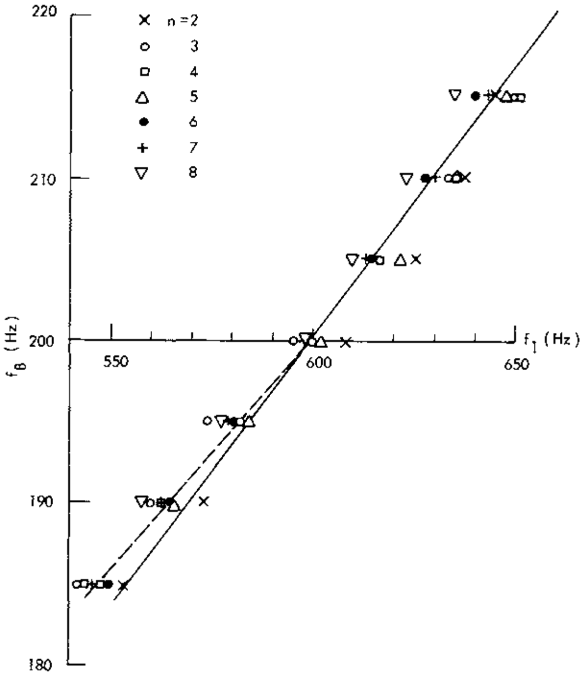


Fig. 19. Results of the A B A X musical-interval matching experiment varying the harmonic numbers for A and B. Stimuli were presented monotically at 40 dB SPL. $f_2 = f_1 + 200$ Hz. Subject A. H.

is drawn through the point $f_1 = 600$ Hz, $f_B = 200$ Hz, with a slope of 1/3, 3 being the harmonic number of the lower partial of X. (The data for downward shifts appear to fit better with a (dashed) line of slope 1/3.5, corresponding to the average harmonic number of X. This detail was not further investigated.)

6.4 EXPERIMENT 17

The results of Experiments 15 and 16 demonstrate a systematic relationship between the stimulus frequencies of an inharmonic tone complex and the missing fundamental of a harmonic complex. We have shown that harmonic tone complexes have a well-defined

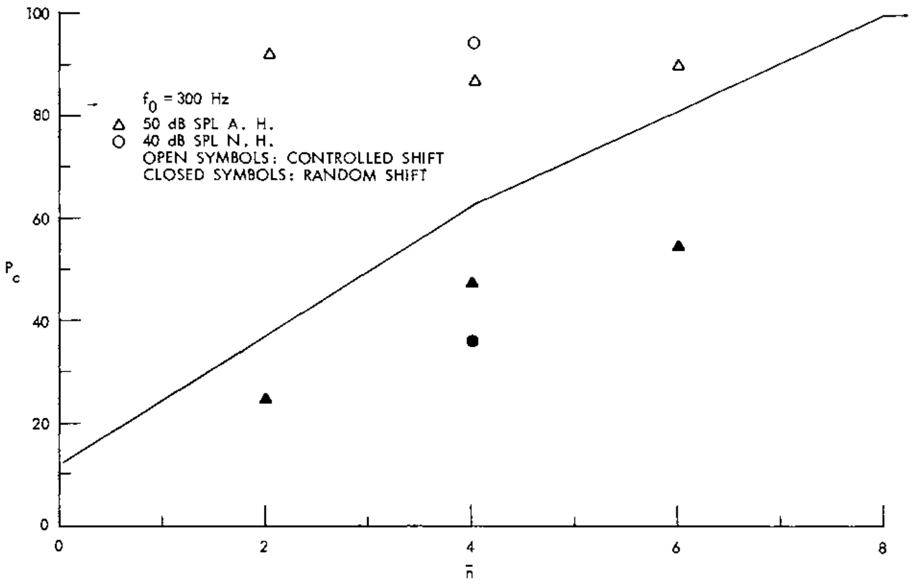


Fig. 20. Effects of inharmonic frequency on interval identification performance of 2 subjects. The graph indicates performance for the random-frequency paradigm predicted by the model explained in Appendix II.

musical pitch; that is, a sequence of harmonic sounds can be perceived as a melody. Another experiment was undertaken to show directly that inharmonic tone complexes can also define a clearly perceptible melody when the empirical relationships described above are employed to control the amount of inharmonic frequency shift.

The basic experimental paradigm was the same as that in Experiments 6-9. Two subjects were asked to identify 8 musical intervals ranging from a major third upward to a major third downward. Notes were represented by two-tone complexes, presented monotically at sound pressure levels of 50 dB for Subject A. H. and 40 dB for Subject N. H. Harmonic numbers were chosen randomly for each note over a range of 3. In addition, an inharmonic frequency shift was introduced for both stimulus frequencies in the following way. For the first note a random frequency shift, positive or negative, was chosen with a maximum of one-fourth of the fundamental frequency; then the empirical relations described in Fig. 18 (slope = 1/n) were used to compute the required frequency shift for the second note that would leave the musical interval unaltered. Figure 20 shows the results of this experiment. It also shows results from a similar interval identification test wherein the amounts of inharmonic frequency shift were chosen randomly and independently for each note. The solid line in this figure represents expected performance with independent shifts calculated on the basis of some simple assumptions which are explained in Appendix II.

6.5 DISCUSSION

To our knowledge, all studies dealing with the pitch of inharmonic tone complexes have employed paradigms of direct comparison; that is, an inharmonic sound whose pitch was being investigated was directly matched to a harmonic reference sound whose pitch was presumed to be known. In our experiments we used a different paradigm: (i) to simulate musical behavior more closely; (ii) because we found that successive notes evoked in the subjects a sense of musical interval and provided a context for the feature of each sound that was being contrasted, and (iii) to minimize the opportunities for behavioral responses that are directly correlated with matching of individual partials.

In spite of these differences in procedure, our experimental findings are in good agreement with those of other investigators. Moreover, our experiments show that similar behavior is obtained with dichotically presented two-tone stimuli. This is further evidence for a common central tracking mechanism that operates similarly for monotic or dichotic stimuli.

No pronounced "second effects" were found in either monotic or dichotic experiments; the results of Experiment 15 can be fitted by straight lines having a slope of $1/n$, as illustrated in Fig. 18. Other investigators have called this the "first effect." The reason why we did not find a clear "second effect" – a consistent deviation from such straight lines – is probably that we used values of n for which both stimulus partials were behaviorally resolvable; when higher values of n are chosen, the central processor may be forced to operate only on those neural signals derived from peripherally resolved aural combination tones that occur only with monotic presentations. The subject found the task very difficult for larger values of n, especially under dichotic stimulus conditions. Perhaps different subjects would reveal different amounts of "second effect."

The data from Experiment 16 can be fitted reasonably well by two lines; deviations can be easily accounted for by the small number of trials associated with each data point. The values of the slopes are consistent with results from Experiment 15 and correspond approximately to the "first effect." The result that the pitch matches do not depend on the harmonic numbers of the (harmonic) reference stimuli is consistent with earlier findings (Sections II and III) that the sensation of a melody played by harmonic tone complexes is not destroyed or altered when different harmonic numbers are chosen from one note to the next.

Experiment 17 shows directly that, as far as musical behavior is concerned, inharmonic two-tone stimuli are equivalent to harmonic stimuli over a considerable range of inharmonic frequency shift. Beyond this range ambiguities may arise (Schouten and others[51]). The results of many other investigations, including our Experiment 15, also indicate that for frequency shifts larger than approximately one-fourth of the fundamental frequency, there are at least two fundamental values of a reference stimulus with which such inharmonic sounds can be associated. When frequency shift does not exceed this range, however, inharmonic two-tone stimuli can be associated unambiguously with a specific musical note. According to our limited data on this matter, the relation between stimulus frequencies and musical note is adequately described by the "first effect," provided that the harmonic numbers and intensities are not too high. This relation is illustrated in Fig. 18, or can be described equivalently as "that subharmonic of the lower frequency component which is closest to the difference frequency" (Walliser[53]). Deviations from these simple rules would be expected on the basis of our own findings (Section III) and those of others (Smoorenburg,[56] Ritsma[55]) that combination tones extend the range of harmonic numbers that are effective in communicating musical pitch. Thus, if the "first effect" or "subharmonic" rules are applied to combination tones, we would predict a "second effect"; that is, the pitch-match data would be described by a line with a reciprocal slope value that is lower than the harmonic number of the lowest stimulus partial.

In Experiment 17 the reason why the subject's performance for independent inharmonic frequency shifts was worse than predicted by a simple model may be that the subject did not use an optimal strategy; not much training was given for this experiment. The slightly less than perfect performance for the case in which frequency shifts were controlled by known empirical relations can easily be explained by the additional subjective difficulty of having to identify musical intervals that did not all begin with the same note, and hence were in a constantly roving key.

The conclusions that we draw from the work described in this section are the following.

1. Despite differences in experimental procedure, there is enough evidence that our studies of fundamental tracking behavior and the work of many other investigators on "periodicity pitch" or "residue pitch" all reflect the same basic phenomenon which is mediated by a central neural mechanism.

2. Inharmonic tone complexes do have a definite musical value, provided that certain conditions are satisfied. This value is approximated by what has been called the "first effect" of inharmonic frequency shift.

3. The frequently reported "second effect" is of secondary importance and is a consequence of the fact that the central mechanism operates only on signals derived from those stimulus partials or aural combination tones that can be resolved behaviorally. Thus the second effect reflects mechanisms that are peripheral to the central processor of musical pitch, and are therefore of no direct relevance in investigating the latter mechanism. The fact that similar bounds on harmonic number apply both when stimulus partials are presented monotically and dichotically suggests that behavioral resolution is dependent on more than just cochlear mechanisms.

VII. SENSITIVITY OF FUNDAMENTAL TRACKING

7.1 INTRODUCTION

Experiments 1-17 have demonstrated some general tendencies and quantitative boundaries in fundamental tracking behavior under several stimulus conditions. On the basis of these findings we have been able to exclude some possible concepts or models of auditory information processing, but we are still far from proposing a unique alternative model. To develop and test such a model, it would be invaluable to have a set of data on the sensitivity of the system to be modeled.

A conventional procedure for studying sensitivity is the two-alternative forced-choice method in which a subject is presented with one of two possible stimuli that differ only in the value of the parameter whose sensitivity is being studied. The subject is instructed to tell which of these two stimuli was presented in a given trial, and his performance is measured as a function of the difference in value of that particular parameter between the two stimuli.

We could imagine such an experiment to measure the auditory system's sensitivity in fundamental tracking. The stimuli would be two-tone complexes, which would be successive random harmonics of the fundamental frequencies f_0 and $f_0 + \Delta f_0$. Correct identification as a function of Δf_0 could then be measured with fundamental frequency, harmonic number, number of harmonics, and intensity as experimental parameters. There are some pitfalls in such a procedure, however. In contrast with the identification experiments the two alternatives in a sensitivity experiment can contain partials that are either nearly coincident or are very nearly related by ratios of small integers. Discrimination of deviations from unisons and small integral ratios is not difficult for trained musicians (Houtsma[62]). Therefore we must be especially cautious in guarding against discrimination merely on the basis of individual partials.

Walliser[53] measured sensitivity for the missing fundamental of highpass-filtered pulse trains and compared the results with sensitivity data for simple tones. He concluded that the former bears little relation to the sensitivity for frequencies equal to the missing fundamental, but can be predicted from the frequency sensitivity data at the individual partials. We could argue that this indicates that the musical pitch sensation corresponding to the fundamental is derived from the sensation of partials, but it seems much simpler to assume that subjects base their decisions directly on the sensation of partials, without paying attention to or perhaps even hearing the musical pitch.

7.2 EXPERIMENT 18

To show that this interpretation is not mere speculation, a control experiment was performed. A subject was presented with a sequence of two-tone complexes having fundamentals of f_0 and $f_0 + \Delta f_0$, or in reverse order. The lower harmonic number was randomly chosen between 2 and 4 for each tone complex, with the condition that no identical harmonic numbers would be chosen in succession in a given trial. Stimuli were presented monotically at a 50-dB sound pressure level. The subject was instructed to try to identify the order of presentation in each trial; feedback was provided after each answer. Percentage correct response was measured in runs of 50 trials for decreasing values of Δf_0, and the value corresponding to a 75% correct response level was defined as the "just noticeable difference" (JND). Figure 21 shows such JND's as a function of f_0. Next, similar runs of 50 trials were taken for values of Δf_0 equal to the JND at the various values for f_0, *but this time with the partials presented sequentially*, as in Experiment 4. The scores for these runs are also indicated in Fig. 21. The fact that over a large range of f_0 the scores are not significantly different from 75% suggests strongly that also in the first part, where both partials were presented simultaneously, decisions were probably based on information coming directly from the partials. The possibility that fundamental tracking exists even when the two partials are presented in time sequence, as in the second part of this experiment, is ruled out by the results of Experiment 4.

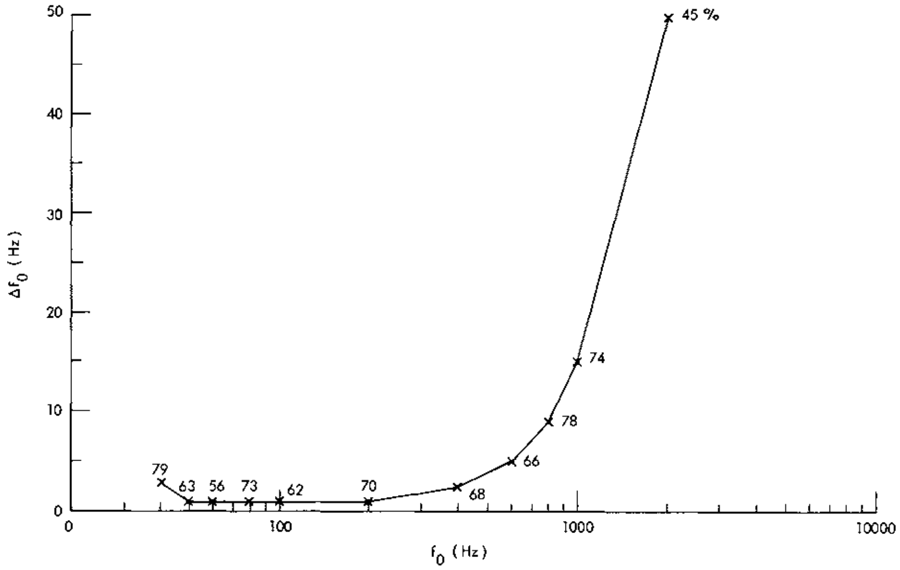


Fig. 21. Just noticeable difference in fundamental frequency measured in a two-interval, two-alternative forced-choice discrimination paradigm. The curve represents a 75% correct performance level for simultaneous presentation of harmonics. The percentages indicate performance levels for time sequential presentation of harmonics.

It seems possible, at least in principle, to extract sensitivity information from the data presented in Section III, given that the limitations in performance are imposed largely by the noisiness of a postulated fundamental percept. It has been shown experimentally that an eight-alternative forced-choice identification paradigm for musical intervals does enable a subject to score perfectly under certain stimulus conditions (see the test experiment with square waves and Experiments 6-10 for low values of $\bar{n}$), which suggests that simple memory limitations are not relevant (Miller[77]). The paradigm effectively prevents a subject from tracking the fundamental merely on the basis of his recognition of individual partials (see Experiment 4). A simple decision model will be explained and employed to transform identification performance data into a single model parameter that may represent the underlying variance of the fundamental percept. At present, this model cannot be justified as being more than a data-reduction scheme. Further work will be required to investigate its theoretical value.

7.3 DECISION MODEL AND DATA REDUCTION

The model is a special case of Thurstone's Comparative Judgment law,[78, 79] and has been developed and applied to intensity perception by Durlach and Braida.[80] After a few more assumptions to make the model more specific, it is defined by the following axioms:

1. Each stimulus presentation, S_i, leads to a particular value of a decision random variable X on a unidimensional continuum, called the <u>decision</u> <u>axis</u>.
2. The observer locates N+1 criteria, $-\infty = C_0 \leqslant C_1 \ldots < C_N = \infty$, on the decision axis.
3. The observer responds R_n only when $C_{n-1} < X < C_n$.
4. The conditional probability density function, $P_i(X/S_i)$, is Gaussian with mean $\mu(S_i) = k \log f_0^i$, where f_0 is the fundamental frequency of the complex stimulus S_i and variance σ_x^2 (independent of i and constant throughout the experiment).
5. The criteria C_i are placed in such a way that the average correct score will be maximized.

This model makes certain predictions which can be tested experimentally. Some tests were performed, but thus far the data are not sufficient to make a firm statement about the applicability of any of the axioms mentioned above. Until enough data are available to confirm the general applicability of the model, it is important to realize that the model should be considered as a transformation on the empirical results of previously described experiments, and be very cautious when attaching any theoretical meaning to the results of the transformation, the model variance σ^2. Two comments with respect to the model assumptions should be made. First, the assumption in axiom 4 that $P(X/S_i)$ is Gaussian is made largely for mathematical convenience. Present work in progress in our Laboratory suggests that the transformation from average percentage correct response to model variance, which is all that we are interested in for present purposes, is rather insensitive to the precise shape of the density function $P(X/S_i)$. The extent to which this would be true when the density functions are multimodal, reflecting the ambiguity often encountered with complex tones (Schouten and others[51]), is still unknown. Second, in axiom 5 the effect of noisy or biased criteria would be a change in the relation between average percentage correct response and model variance. As long as the average bias and criterion noise do not change too much with signal parameters like fundamental frequency or harmonic number, the result will be that all computed values for the model variance will be affected equally, thereby changing only the absolute levels, but leaving relative behavior unchanged.

This model is illustrated in Fig. 22. All densities are Gaussian and have equal variance σ^2. Since the means are proportional to the log of the fundamental frequencies, and only the second note, f_0', in the sequence f_0, f_0' contains relevant information on each given trial, and these eight notes have the values illustrated in Fig. 4a, it follows

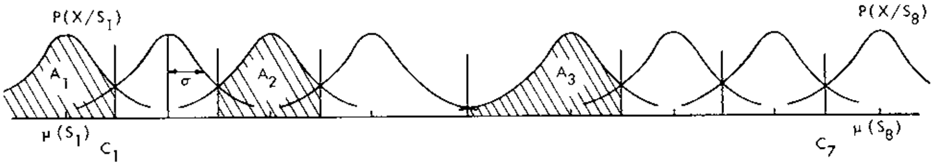


Fig. 22. Illustrating the decision model.

that the distances between successive means are equal, except for a doubled distance between $\mu(S_4)$ and $\mu(S_5)$ (a full-tone frequency ratio is exactly the square of a semitone ratio in the equally tempered scale). It can also be shown that in order to maximize the average correct score, criteria should be placed at the points where successive densities intersect, or equivalently, halfway between adjacent means.

The average percentage correct response, P_c, is then given by

$$P_c = \Pr(R_1/S_1)\Pr(S_1) + \Pr(R_2/S_2)\Pr(S_2) + \ldots + \Pr(R_8/S_8)\Pr(S_8). \tag{1}$$

Since all stimuli are equally likely to occur in a given trial, we can write

$$P_c = \int_{-\infty}^{C_1} P(X/S_1)\,dX + \frac{1}{8}\int_{C_1}^{C_2} P(X/S_2)\,dX + \ldots + \frac{1}{8}\int_{C_7}^{\infty} P(X/S_8)\,dX. \tag{2}$$

By referring to Fig. 22, this expression can be written

$$P_c = \frac{1}{4}A_1 + \frac{1}{2}A_2 + \frac{1}{4}A_3, \tag{3}$$

where

$$A_1 = \int_{-\infty}^{C_1} P\left(\frac{x}{S_1}\right) dX = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\frac{\mu(S_1)+\mu(S_2)}{2}} \exp\left[\frac{1}{2}\left(\frac{x-\mu(S_1)}{\sigma}\right)^2\right] dX$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\mu(S_2)-\mu(S_1)}{2\sigma}} e^{-\frac{1}{2}Y^2} \, dY \triangleq \phi\left(\frac{d'}{2}\right), \qquad (4)$$

with

$$d' = \frac{\mu(S_2) - \mu(S_1)}{\sigma} \quad \text{and} \quad \phi(X) = \int_{-\infty}^{X} \frac{1}{2\pi} e^{-\frac{1}{2}Y^2} \, dY.$$

$$A_2 = 1 - 2(1-\phi(d'/2)). \qquad (5)$$

$$A_3 = 1 - (1-\phi(d'/2)) - (1-\phi(d')). \qquad (6)$$

From (3) and (6) it follows that

$$P_c = 1.5\phi(d'/2) + 0.25\phi(d') - 0.75. \qquad (7)$$

This function is illustrated in Fig. 23; it transforms the experimentally observed performance level into a model sensitivity index, d', from which the model variance can be computed:

$$d' = \frac{\mu(S_{i+1}) - \mu(S_i)}{\sigma} = \frac{k \log\left(f_0^{i+1}/f_0^i\right)}{\sigma} \qquad \text{for } i \neq 4. \qquad (8)$$

By assuming that the frequency ratio f_0^{i+1}/f_0^i is an equally tempered semitone for $i \neq 4$, instead of the natural intervals that were actually used, we accepted a negligible error as the price for greatly simplified computations. Taking the logarithm to the base 2, and choosing k = 1200, the standard deviation, σ, will be expressed in cents, and is given by

$$\sigma = \frac{1200 \log_2 \sqrt[12]{2}}{d'} = \frac{100}{d'}. \qquad (9)$$

For every point on the performance graphs in Section III we can now compute the corresponding model variance. Instead of doing this for each subject and each individual experiment, we assumed that the results of the low-intensity monotic and dichotic and of the 50-dB SPL dichotic experiments could be considered essentially identical. Therefore "average" performance contours were plotted for each subject, averaged over Experiments 6, 7, and 9, by taking the average $\bar{n}$ for each given

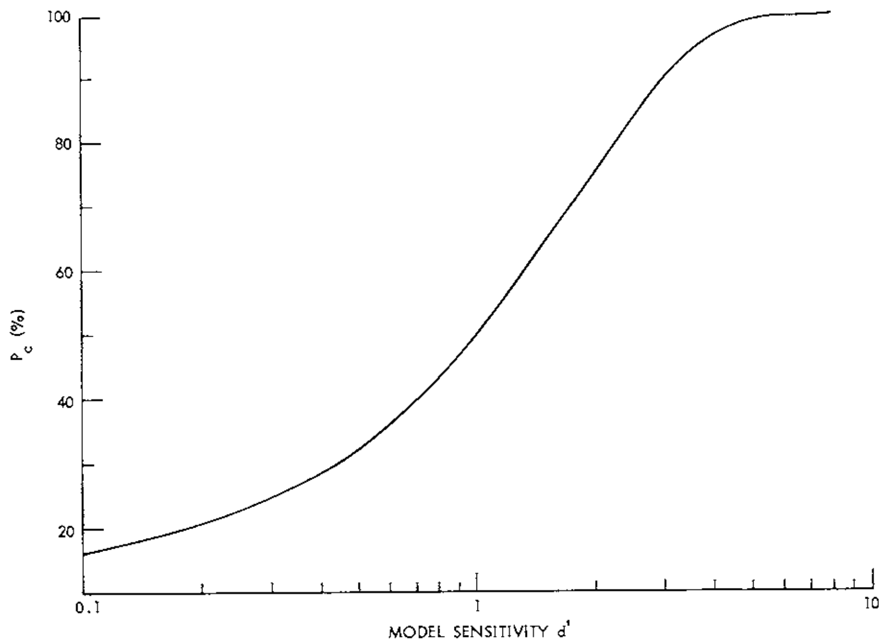


Fig. 23. Relation between model sensitivity, d', and average performance, P_c, in an eight-alternative identification paradigm as described in Section III.

combination of f_0 and performance level (see Fig. 24). From these contours the function $\sigma(n, f_0)$ was computed for each of three subjects, using the relations (7) and (9). These functions are plotted in Figs. 25 and 26 for constant $\bar{n}$ (horizontal cuts through Fig. 24) and constant f_0 (vertical cuts) in logarithmic coordinates, with the product $\bar{n}f_0$ as independent variable. They represent the model standard deviation in cents at the fundamental frequency f_0 as a function of the average stimulus frequency. Figure 26 suggests that the standard deviation can be described by the compact expression

$$\sigma(n, f_0) = H[G(f_0)\bar{n}f_0], \tag{10}$$

where $G(f_0)$ is the horizontal intercept function, and $H[X]$ the characteristic curve.

$H[X]$ was determined empirically by overlaying the plots of Fig. 26 for all fundamental frequencies and all subjects, and fitting a curve by eye. The result is shown in Fig. 27 and is well described by two straight lines with slopes 2.3 and 5.7, respectively.

The inverses of the horizontal intercepts, $I(f_0)$ = reciprocal of abscissa for σ = 100 and fundamental f_0 in Fig. 26, are plotted in Fig. 28 for each subject. Since $H(1) = 100$ by definition (Fig. 27), we require $G(f_0) = I(f_0)$. $I(f_0)$ can be fitted with the general form

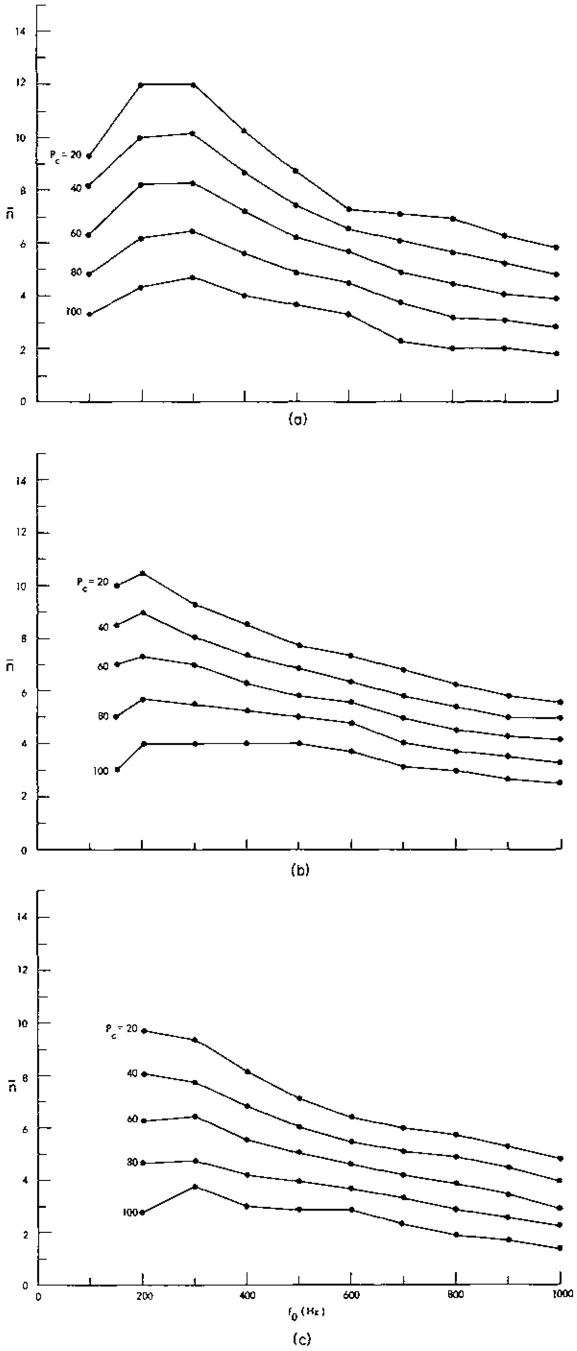


Fig. 24. Average performance contours. (a) Subject N. H. (b) Subject A. H. (c) Subject S. W.

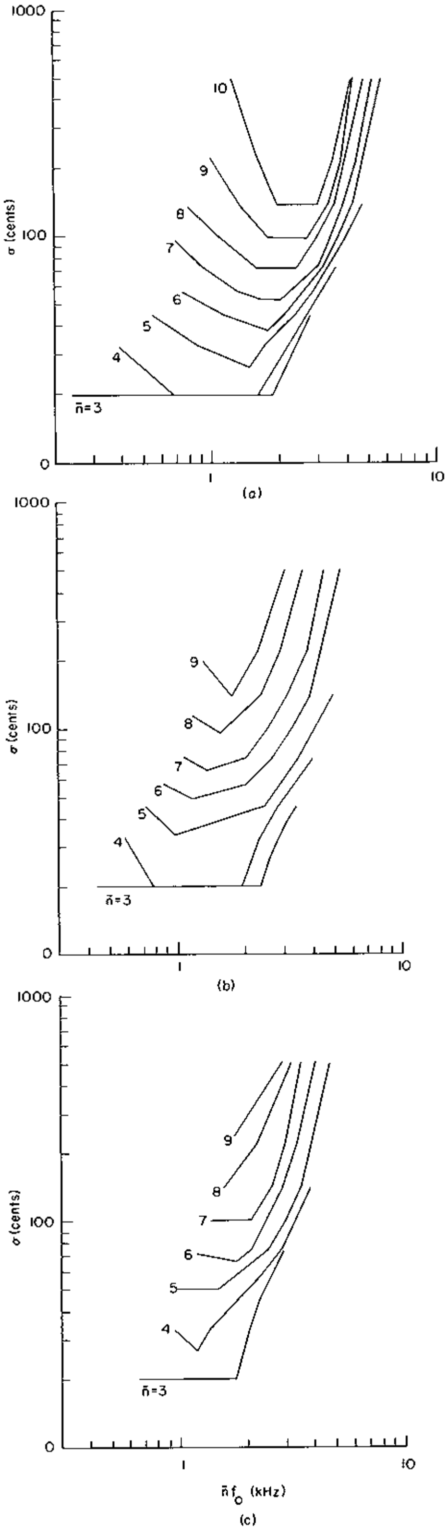


Fig. 25. Model standard deviation vs average stimulus frequency for constant $\bar{n}$. (a) Subject N.H. (b) Subject A.H. (c) Subject S.W.

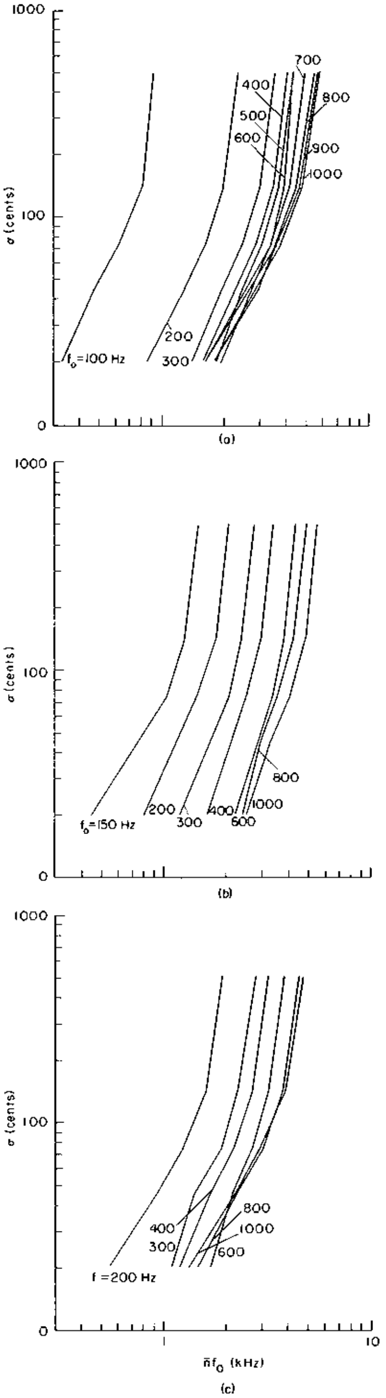


Fig. 26. Model standard deviation vs average stimulus frequency for constant f_0. (a) Subject N.H. (b) Subject A.H. (c) Subject S.W.

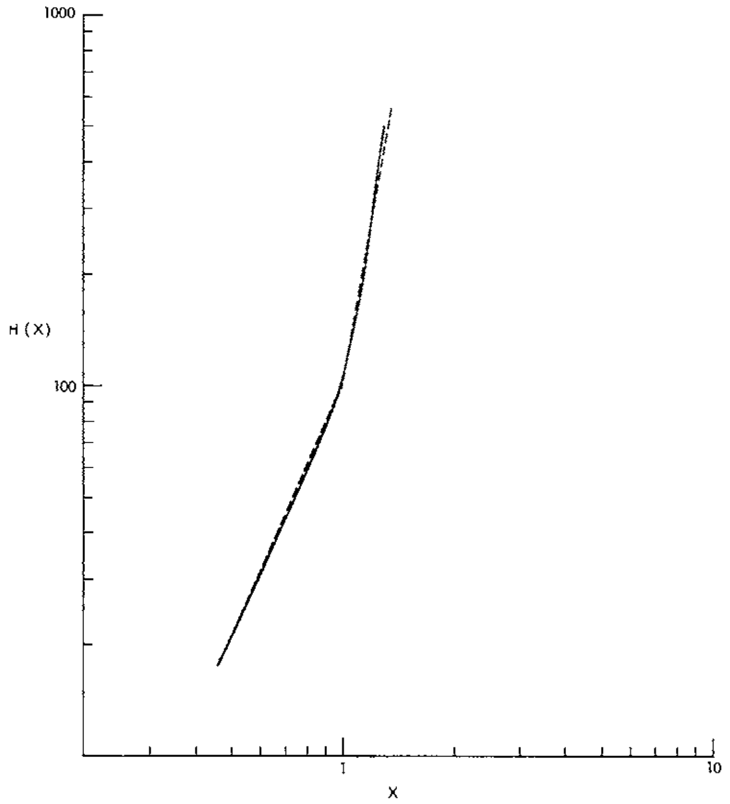


Fig. 27. Average representation of functions shown in Fig. 26. Dashed curve is a two-piece linear approximation.

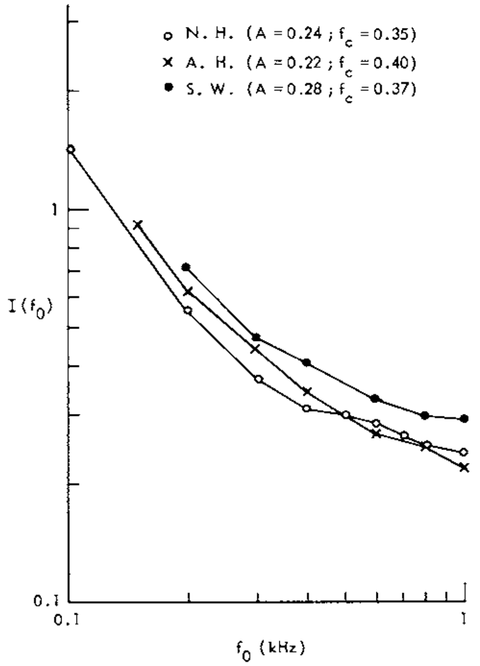


Fig. 28.
Horizontal intercept function $I(f_0)$ for 3 subjects taken from Fig. 26.

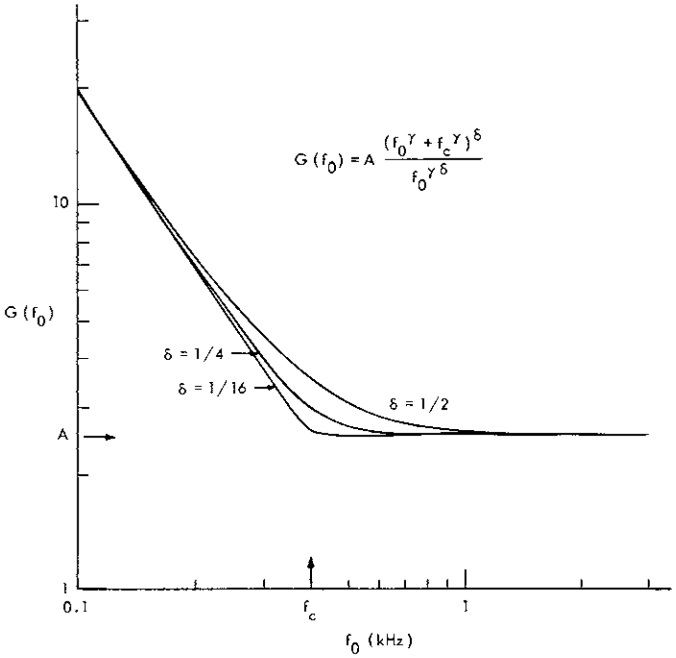


Fig. 29.
The function $G(f_0)$ for $\gamma\delta = 1.5$ and several values of δ.

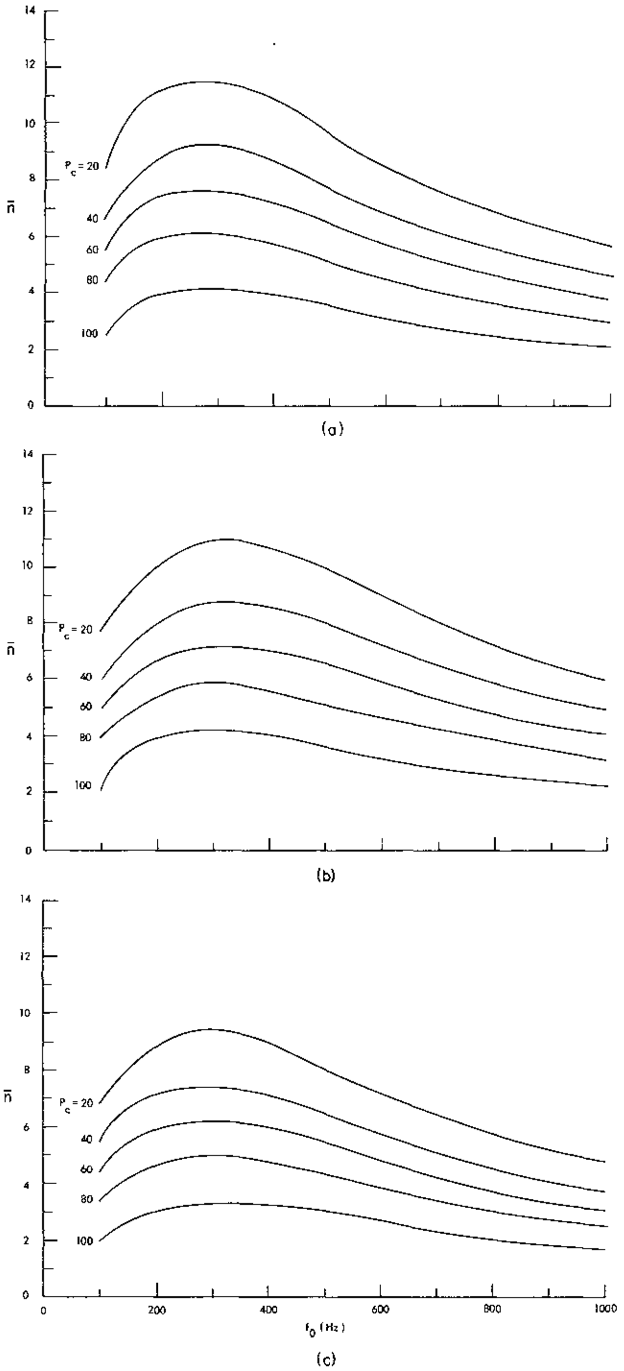


Fig. 30. Equal performance contours determined by expression (12). (a) Subject N. H. (b) Subject A. H. (c) Subject S. W.

$$G(f_0) = \frac{A\left(f_0^3+f_c^3\right)^{0.5}}{f_0^{1.5}} \tag{11}$$

as shown in Fig. 29. Particular values for the parameters A and f_c are indicated in Fig. 28 for each subject.

Summarizing, we now have the following compact description of σ, the model standard deviation:

$$\sigma(n, f_0) = H\left[A\bar{n}f_0 \frac{\left(f_0^3+f_c^3\right)^{0.5}}{f_0^{1.5}}\right] = H\left[A\bar{n}\left(\frac{f_0^3+f_c^3}{f_0}\right)^{0.5}\right], \tag{12}$$

where

$$H[X] = 100\ X^{2.3} \qquad \text{for } X < 1$$

$$= 100\ X^{5.7} \qquad \text{for } X \geqslant 1$$

and f_c and f_0 are expressed in kHz.

We can test how much information was destroyed by this data-reduction procedure by computing P_c for various combinations of $\bar{n}$ and f_0, using (7), (9), and (12), and comparing the results directly with the experimental results of Section III. Figure 30 shows that the original data can be reproduced quite well from expression (12), which ensures that (12) gives a reasonable description of the data.

By taking derivatives of (12) with respect to $\bar{n}$ and f_0, we can show that $\sigma(n, f_0)$ has a minimum, and hence P_c is maximum for $\bar{n} = 0$ and $f_0 = f_c/\sqrt[3]{2}$.

7.4 RELATION TO FREQUENCY SENSITIVITY FOR SIMPLE TONES

Now we shall consider whether the model variance which was computed from results of experiments with complex tones bears any relation to the sensation variance for simple tones. If f_0 is the missing fundamental of a complex sound, and f_1 and f_2 are the frequencies of two successive partials, and if we assume that each partial leads to random decision variables, X_1 and X_2, respectively, which are independent and have Gaussian densities with means proportional to frequency (not log frequency), then the least mean-square error estimate of the fundamental f_0 from the observed variables X_1 and X_2 is approximately given by

$$X_0 = X_2 - X_1.$$

[Note: We have neglected the improvement in the estimate that can be derived in principle from the a priori information that $\overline{X}_2$ and $\overline{X}_1$ are successive harmonics.]

The variances of X_1 and X_2 are known from several experiments on differential sensitivity for frequency (Shower and Biddulph,[81] Koester,[82] Harris[83]). Approximating these data by a constant Weber fraction, $\Delta f/f = C$, and noticing that for a two-alternative, two-interval discrimination paradigm $\Delta f_1 = \sigma_{X_1}$ and $\Delta f_2 = \sigma_{X_2}$ when discrimination performance is 75% correct, we can compute the variance of X_0 and hence predict performance for Experiments 6-10, under the assumption that the decision model applies. For any combination of $\bar{n}$ and f_0 on a frequency decision axis we have

$$\sigma_{X_0} = \sqrt{(Cnf_0)^2 + (C(n+1)f_0)^2} \text{ Hz.} \tag{13}$$

Converting back to a log-frequency axis, and expressing the standard deviation in cents, we have

$$\sigma_{Y_0} = 1200 \log_2 \left[\frac{f_0 + \sigma_{X_0}}{f_0}\right] \text{ cents.} \tag{14}$$

From expression (14) the expected performance can be computed by using expressions (7) and (9).

Figure 31 shows the computed standard deviation at f_0, expressed in cents, as a function of the product $\bar{n}f_0$ for several values of $\bar{n}$. A Weber fraction of 0.002 was chosen, which is consistent with the available frequency discrimination data, at least for frequencies somewhere between 0.5 kHz and 5 kHz. Figure 31 should be compared with Fig. 25. A similar plot for constant f_0 is shown in Fig. 32, which should be compared with Fig. 26. Finally, predicted equal performance contours are plotted in Fig. 33, which should be compared with Fig. 24.

From these comparisons it is evident that the predicted variances are generally much smaller than those actually measured, and that both behave quite differently as a function of the partial frequency $\bar{n}f_0$. Equivalently, predicted performance generally is much better than was actually observed. There are several possible reasons for this discrepancy. First, as we have mentioned, it is quite possible that the standard deviations presented in Figs. 25 and 26 are merely arbitrary transformations on performance data and do not have a psychophysical meaning, the general applicability of the proposed decision model being the crucial issue. In this case there is no reason why they should be consistent with variances predicted from optimal processing of simple tones. Second, if it can be shown that the decision model is generally applicable and that the computed model variance can be regarded as representing the variance for the missing fundamental, it is still possible that the process of deriving the fundamental from successive partials is not optimal. In this case the measured variance will always be larger than that obtained by optimal processing, which would definitely be consistent with the findings presented here.

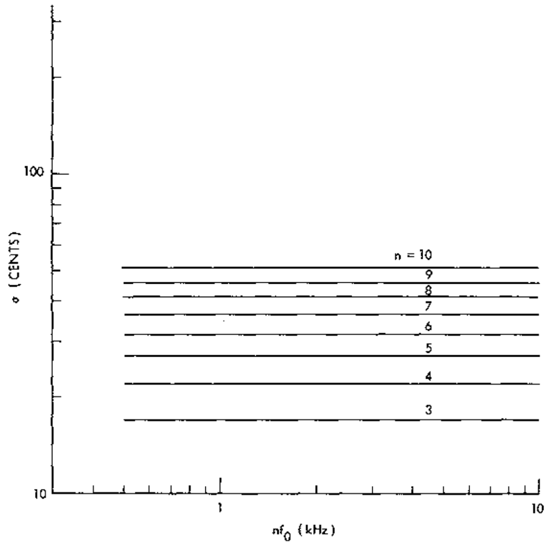


Fig. 31.

Standard deviation of the random variable $Y = X_1 - X_2$ plotted for constant n. The standard deviations of X_1 and X_2 are determined by the Weber fraction $\sigma_{X_1}/f_1 = \sigma_{X_2}/f_2 = 0.002$.

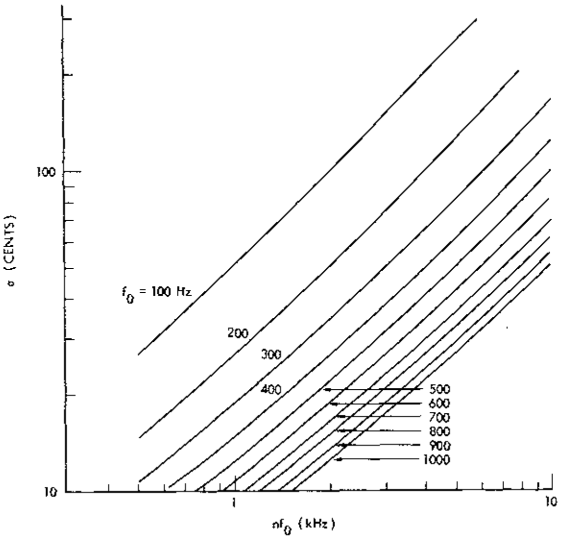


Fig. 32.

Standard deviation of the random variable $Y = X_1 - X_2$, plotted for constant f_0.

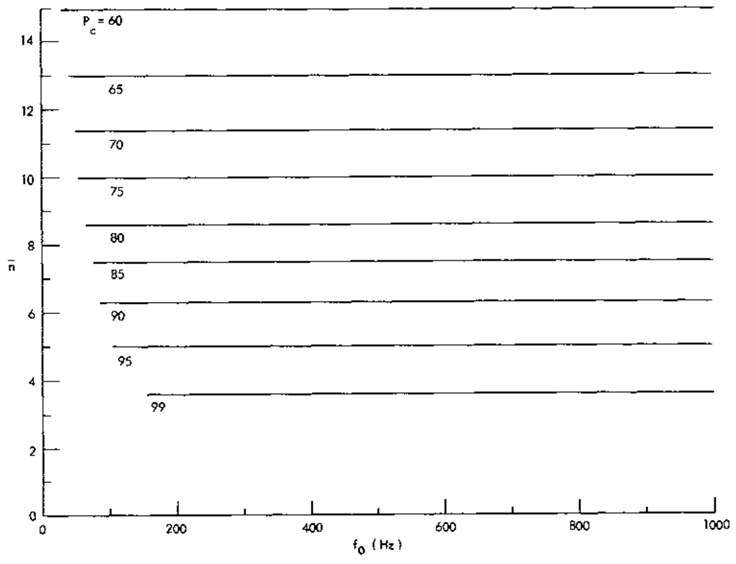


Fig. 33. Theoretical performance contours obtained by optimal processing of simple tone percepts, under the assumption of a Weber fraction of 0.002.

7.5 SUMMARY

Preliminary steps were taken toward a quantitative characterization of the limits in the precision of the mechanism responsible for fundamental tracking. Experimental evidence has been given that conventional two-alternative discrimination experiments may allow the subject to operate in a mode that is irrelevant to fundamental tracking. A preliminary attempt was made to extract information on the precision, or sensitivity, of the fundamental tracking mechanism by assuming that the performance in identifying musical intervals is limited only by the sensitivity to single notes. Because of the number of untested assumptions that were made, we emphasized that the derived sensitivity measure may have little theoretical meaning beyond being a transformation of our one set of nonredundant data; that is, the decision model used for this transformation may not generalize to other stimulus conditions. Finally, we examined the question of whether the limits on musical-interval identification could be attributed to limitations in discriminating the frequency of simple tones, and found that this is not possible because predicted performance is superior to actual performance and both bear different relations to the stimulus parameters.

VIII. CONCLUSION

The object of our study has been the perception of musical intervals. It is evident, however, that the object is closely related to, and, as a matter of fact, is just a different approach to, what has often been described as the perception of musical pitch. The word "pitch" was used in this report in the context of musical operations, in a manner similar to that used by musicians all the time without apparently causing much confusion. Indeed, the design of each of our experiments incorporated an operational conception of the words "melody" and "pitch":

"A melody sensation is the subjective correlate of a sequence of musical sounds corresponding to notes on a musical scale; its presence or absence is determined by the consistent recognition of such a sequence" (see Appendix III).

"Musical pitch is the subjective correlate of each of the musical sounds in such a sequence" (see Appendix III).

Although we regard melody sensation as a sequence of musical pitch sensations, the question of the pitch of an individual sound never arose. While it is true that a musician often refers to the pitch of this or that note, he always does this in the context of other notes, for example, a melody. If an individual sound is taken out of its musical context, its pitch, that is, the note that one would normally associate it with in a musical context, is often ambiguous. For example, more than one note value can often be associated with one complex sound (Schouten, Ritsma, and Cardozo[51]). Furthermore, a subject can attend to different perceived aspects of a sound, and this ability is strongly influenced by learning (Cross and Lane[66]). Psychophysicists who have spent long hours attending to the sensations evoked by individual spectral components of a sound may be particularly hard pressed to ascribe one particular musical pitch to an isolated tone complex as a whole. Helmholtz remarks how difficult it is for musically trained observers to perform the introspective spectrum analysis that he so admirably trained himself to do (Helmholtz[9]). The converse would appear to apply to the ability to hear "musical pitch." On the other hand, one's attention to "spectral cues" appears to be distracted when listening to a sequence of musical sounds. Finally, the ability to assign notes to single musical sounds depends greatly upon special training and perhaps rare inherited talent (Bachem,[84] Brady[85]).

Questions concerning the ambiguity in the meaning of pitch can be avoided by investigating the subject's ability to perform tasks that are musically meaningful, such as identification of melodies or simple intervals. Performance in these tasks depends on the subject's ability to "hear" the musical value or note that each sound represents. Moreover, it is clear that what is being tested is the ability to hear melodies or intervals; this does not require from the subject any conscious, definitive appreciation of the musical pitch of each note.

The use of interval recognition to study the pitch of musical sounds is not a new idea. Writing on the sensation of tone, Mach[86] remarked: "The ability to pick out and recognize intervals is the first thing required of the student of music who is desirous of becoming thoroughly familiar with his subject." The key assumption here is that a certain pitch sensation for each individual sound is a necessary condition for recognizing a melody. We have shown that in simple comparison paradigms like Experiment 18 a subject can apparently discriminate between two fundamentals by using the pitch sensations of single partials which have no fixed relation to the fundamental. In a task which is less artificial and resembles musical behavior more closely, like the experiments in Sections II and III, pitch sensations of individual partials are not sufficient to provide the subject with adequate information about a sequence of fundamentals. The subject appears to require a single sensation, musical pitch, for each successive sound which he can directly associate with a note of the melody.

Our operational definition of musical pitch in relation to melody necessarily limited this project to the study of "relative pitch." The phenomenon of "absolute pitch," which is the ability claimed by certain people for recognizing and identifying the pitch of an individual sound and placing it on an absolute note scale without the aid of a reference note, was not considered in this study (Bachem,[84] Brady,[85] and for an excellent review of the phenomenon of absolute pitch see Ward[87]). Clearly, absolute pitch is not a necessary condition for melody recognition. Moreover, it seems that identification experiments for single notes, employing a subject with genuine absolute pitch, would have created the same problems as matching experiments of static sounds, namely that the subject can often choose to pay attention to one of several features in a sound, not necessarily the one that has the dominant musical relevance.

In Section I we stressed the importance of the auditory system's ability to track simultaneous melodies in listening to music. Helmholtz was originally criticized for putting too much emphasis on harmony and not enough on melody in his theory of music. His response stresses the importance of melody.

> "As to my theory of consonance, I must claim it to be a mere systematisation of observed facts (with the exception of the functions of the cochlea of the ear, which is moreover an hypothesis that may be entirely dispensed with). But I consider it a mistake to make the theory of consonance the essential foundation of the theory of music, and I had thought that this opinion was clearly enough expressed in my book. The essential basis of music is melody. Harmony has become to Western Europeans during the last three centuries an essential, and, to our present taste, indispensable means of strengthening melodic relations, but finely developed music existed for thousands of years and still exists in ultra-European nations, without any harmony at all."[88]

The work reported here suggests that Helmholtz' conceptions of consonance and dissonance require re-evaluation. When several tones in a complex sound bear a simple harmonic relation to a particular frequency, they will strengthen the sensation of that frequency, so that the listener can associate a note of that frequency with the whole complex. This enhancement of a fundamental pitch is weakened when simple harmonic relations are disturbed, either by the exclusive presence of partials with large harmonic numbers or by the presence of inharmonic partials. Although we investigated this phenomenon under laboratory conditions and not in a concert hall, using sounds that are very simple compared with sounds produced by conventional musical instruments, we still believe that our experiments simulated musical behavior sufficiently well to ensure that the reported phenomenon plays a significant role in the practice of music.

From a converse point of view, Helmholtz's statement implies that dissonance, which he conceived as being due to beating partials, weakens melodic relations. Our experiments have indicated that indeed melodies played by upper partials only become less recognizable with increasing harmonic number. This finding appears to be consistent with Helmholtz' explanation of dissonance as peripheral interference of stimulus tones, creating beats. Nevertheless, our findings are inconsistent with Helmholtz' theory because similar failures in melody recognition were found for monotically, as well as dichotically, presented partials; in the latter case peripheral (cochlear) interference cannot play a role. They are much more consistent with the following observation by Mach.

> "A tuning fork held before one ear is very feebly heard by the other ear. If two slightly discordant, beating tuning forks are held in front of the same ear, the beats are very distinct. But if one of the forks is placed before one ear, and the other before the other, the beats will be greatly weakened. Two forks of harmonic interval always sound slightly rougher before one ear. But the character of the harmony is preserved when one is placed before one ear. The discord also remains quite perceptable in this experiment. Harmony and discord are, however, not determined by beats alone."[86]

Indeed, the close correlation which was established in our experiments between melody recognition and behavioral resolution of partial frequencies suggests a close relation between consonance and melodic clarity, or, conversely, between dissonance and lack of melodic clarity. The behavioral similarity for monotic and dichotic stimulus conditions indicates, however, that dissonance cannot be contributed to beats alone, but must have a more central cause.

Introspectively it seems that explaining dissonance by beats is somewhat like fitting a sensation to a theory rather than the other way around. If we listen to an orchestral sound comprising many notes that do not have any simple harmonic relations, and are played by different instruments with complex spectra, there are usually so many tones beating at different rates that actually no beats at all can be perceived; nevertheless, the dissonant quality of such a sound is very pronounced. It seems to do more justice to musical experience to describe a consonant sound as a sound whose components converge to one point; they all enhance the perception of the fundamental note, which is often the bass note. A dissonant sound diverges; there is not any one note whose perception is being strengthened by other components, which makes the sound "atonal" or "pitchless," or, in an operational sense, makes it difficult for the listener to associate any particular musical note with the sound. This introspective observation lends further support to the musical relevance and importance of the phenomenon that was investigated in this study. Thus partials that are harmonic, successive, and of low order are efficient conveyers of musical pitch; therefore, periodic sounds of low harmonic number are relatively more consonant than those of higher harmonic number.

Perhaps the most important finding of this investigation is that the phenomenon of fundamental tracking must be mediated by a central neural mechanism. The results of dichotic experiments have shown conclusively that all cochlear mechanisms, such as two-tone inhibition, beats, combination tones, and neural synchrony to the difference tone are either irrelevant or have secondary importance. Specifically, the "residue theory," which is entirely based on cochlear interaction, is definitely insufficient, and probably completely irrelevant in explaining fundamental tracking phenomena. The word "residue" in this report is always in quotation marks, since we feel that our experiments do not just contradict its definition (Schouten and others[51]) but the very word itself.

Very little was said specifically about the neural mechanism that mediates fundamental tracking. Its close connection with behavioral frequency resolution suggests that neural signals that allow one to recognize simple tones and the missing fundamental of harmonic complex tones are common in early stages of processing; whether this means that the sensation of such a missing fundamental is mediated by the sensation of partials (Terhardt,[89] Walliser[53]) is uncertain.

Whatever kind of neural mechanism is postulated, it will operate on neural signals derived from peripherally resolved tones. It is known that information of the frequency of such tones is preserved after the neural transformation in the form of the place of active nerve fibers and, at least for lower frequencies, in the temporal firing patterns of individual fibers (Kiang[21]). Although in this study certain experimental constraints were found which may turn out to be relevant, such as the mechanism's insensitivity to the relative phase of partials and its behavior for inharmonic partials, it seems that all of these constraints could conceivably be met by mechanisms based on either time or place information. An important simplification for theoretical models is afforded by the empirical finding that fundamental tracking for dichotically and monotically presented two-tone stimuli were shown to be essentially identical. Thus we have a stimulus situation for which descriptive models of the stochastic transformation from stimulus tone to nerve signal have been developed and tested (Siebert,[90] Gray,[91] Evans,[92] Colburn[93]). For any particular mechanism that we assume to operate on those nerve signals, whether it is a time or a place mechanism, we can in principle compute a sensitivity because the stochastic nature of the input to such a mechanism is known. Empirical sensitivity data are thought to be very important in enabling the testing of any postulated model for fundamental extraction.

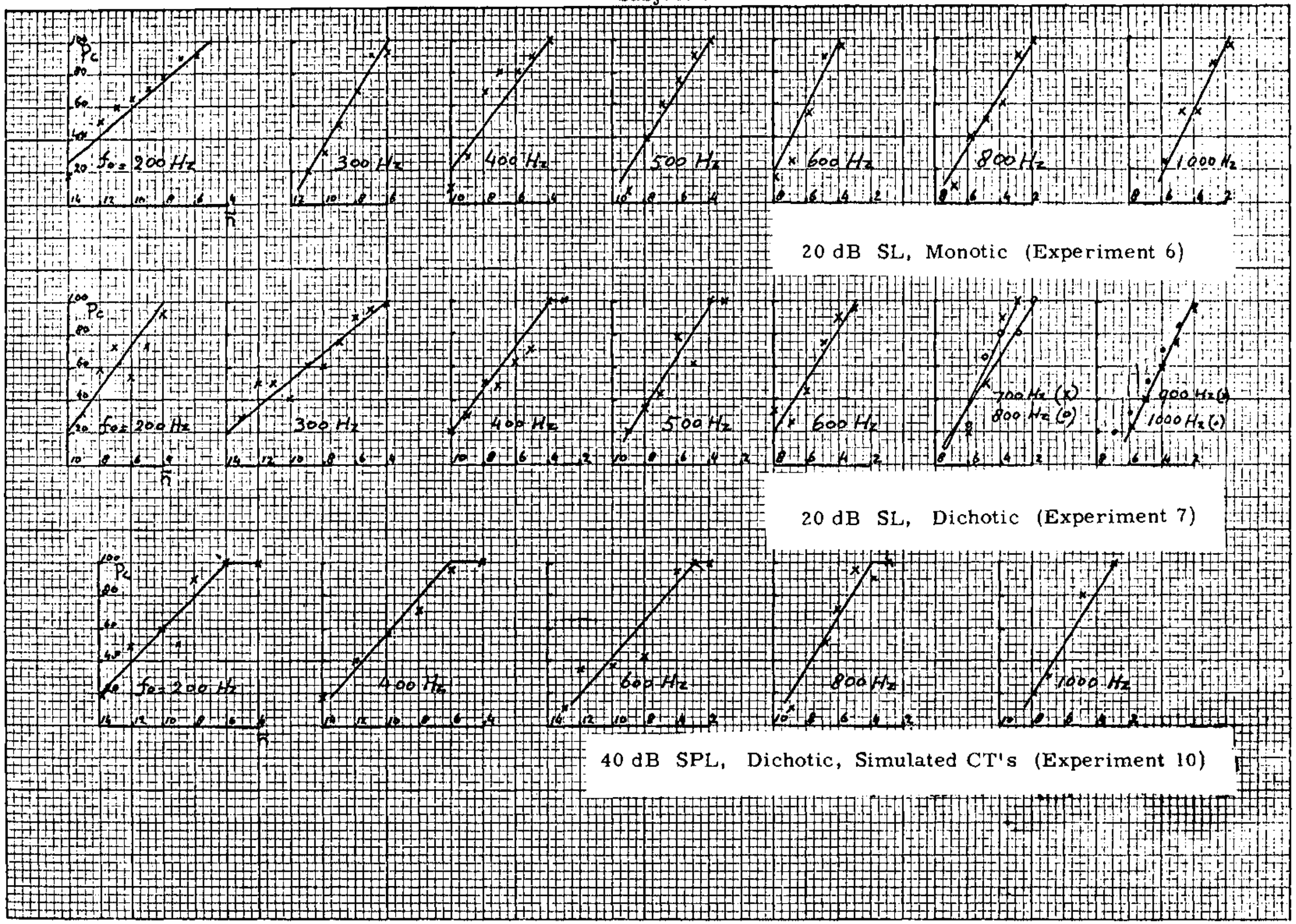
The significance of the sensitivity data presented in Section VII depends greatly on the general applicability of the decision model. More experiments will have to be undertaken to find out specifically which of the model assumptions are tenable and which are not, and we shall have to compute how sensitive the data-reduction procedure is to those assumptions that are not supported by experimental evidence.

Finally, the results of the experiments discussed in this report have an important implication for future physiological research. In the past, efforts have been made to find synchrony effects in the firing patterns of nerve units in the peripheral auditory system to the periodicity or missing fundamental of complex input stimuli. It is clear from the data presented in this report that if any relevant physiological correlates of the missing fundamental could be found, we would have to look in a more central part of the auditory nervous system rather than in the periphery. Activity synchronous with difference tone or missing fundamental, which has been found to exist in the peripheral auditory system (Kiang,[21] Rose and others[94]) is definitely insufficient to explain, and is probably of little relevance to the phenomenon of fundamental tracking.

APPENDIX I

Psychometric Functions for Experiments 6-11 in Sections III and IV

Subject N. H.

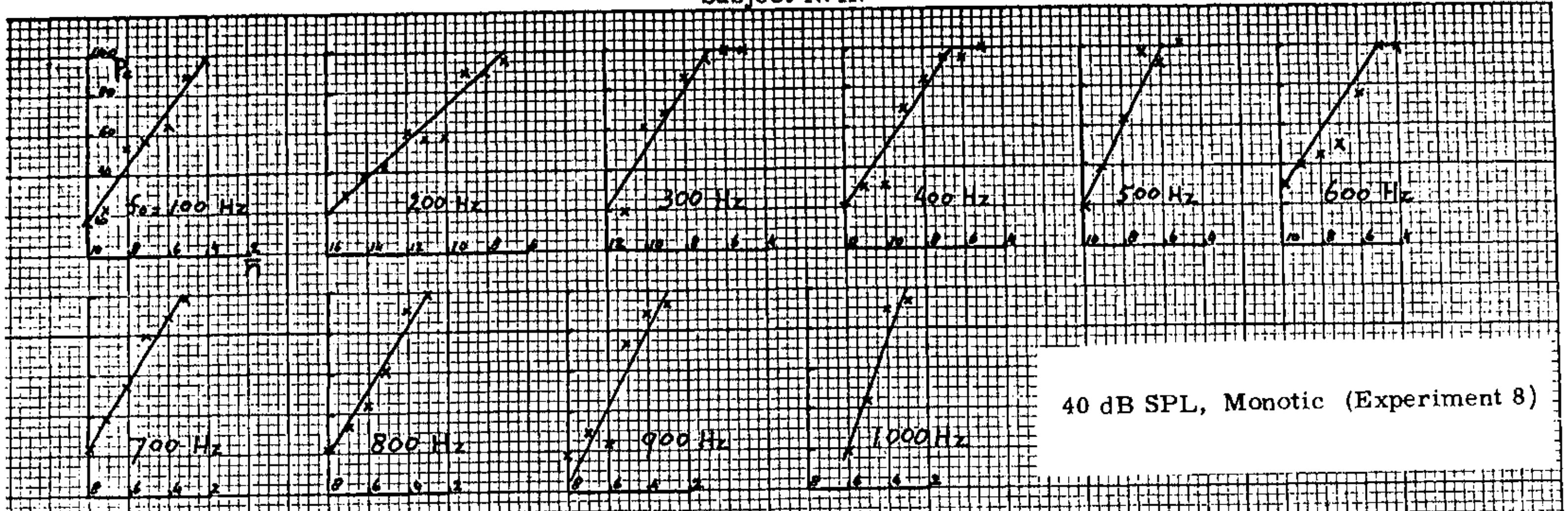


20 dB SL, Monotic (Experiment 6)

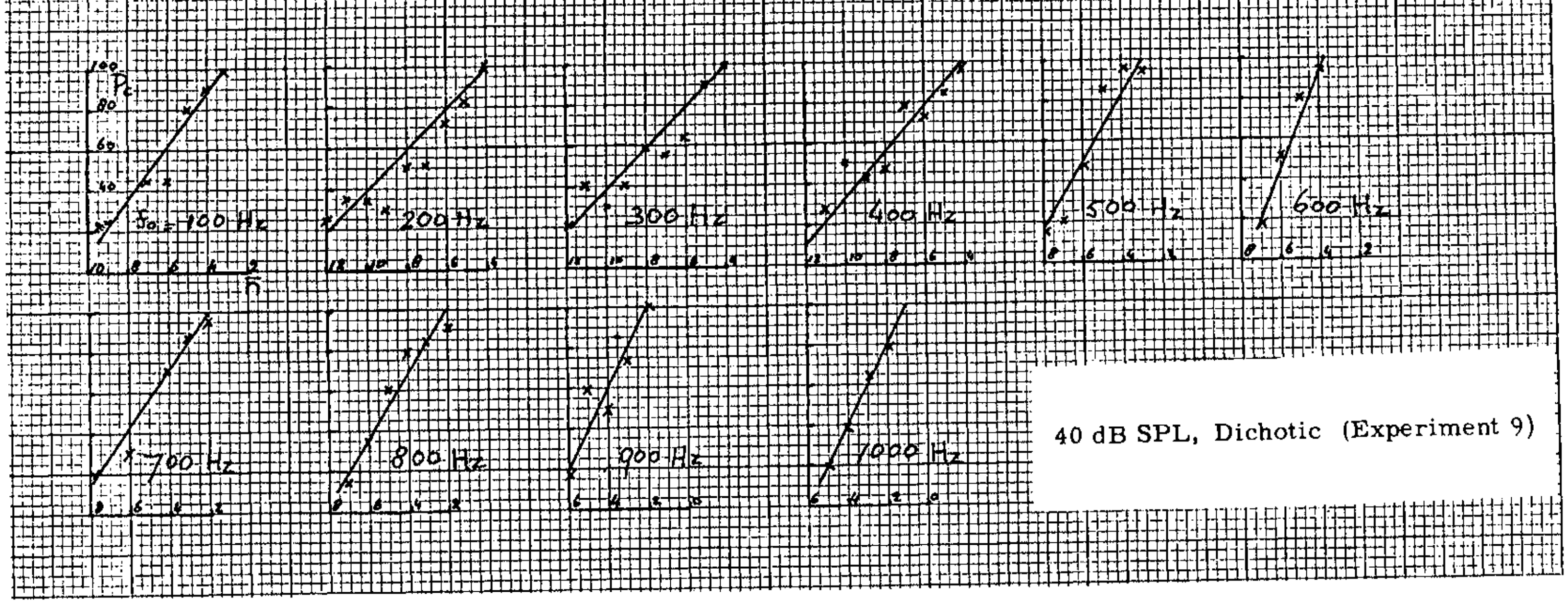
20 dB SL, Dichotic (Experiment 7)

40 dB SPL, Dichotic, Simulated CT's (Experiment 10)

Subject N. H.

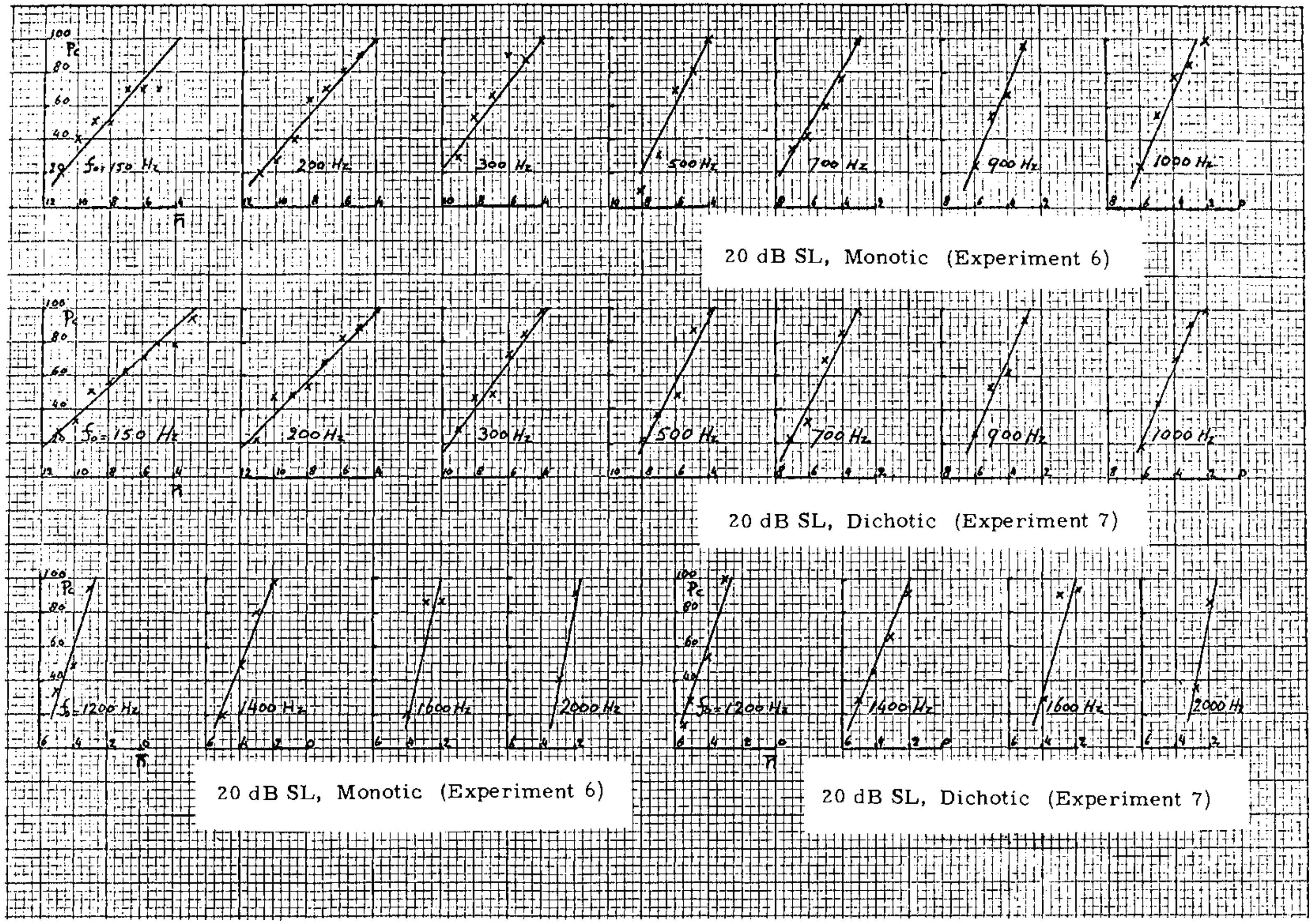


40 dB SPL, Monotic (Experiment 8)



40 dB SPL, Dichotic (Experiment 9)

Subject A. H.



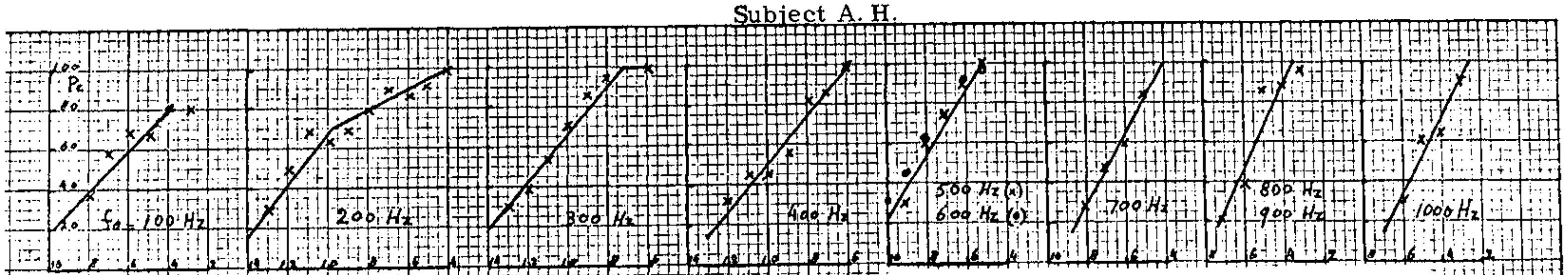
20 dB SL, Monotic (Experiment 6)

20 dB SL, Dichotic (Experiment 7)

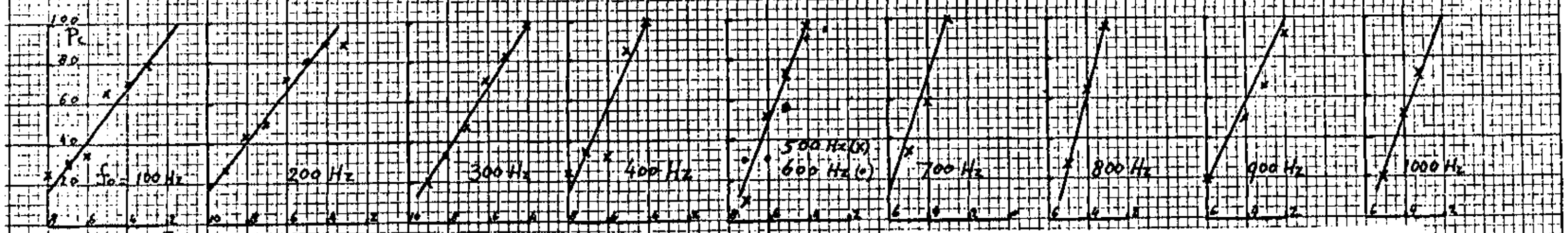
20 dB SL, Monotic (Experiment 6)

20 dB SL, Dichotic (Experiment 7)

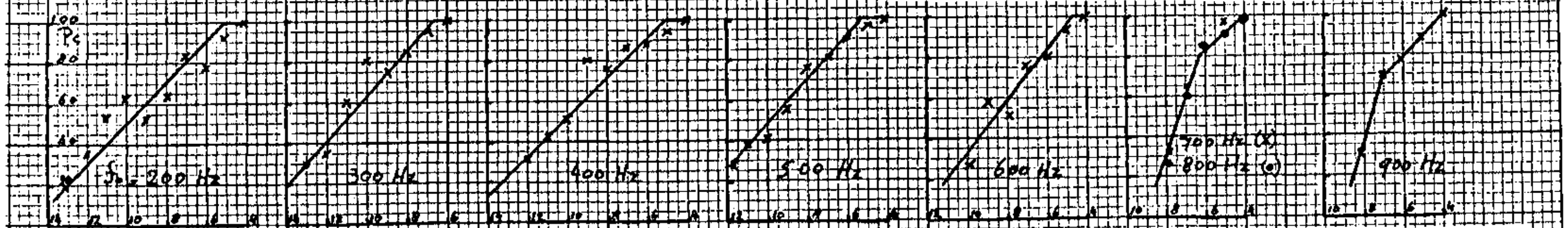
Subject A. H.



50 dB SPL, Monotic (Experiment 8)

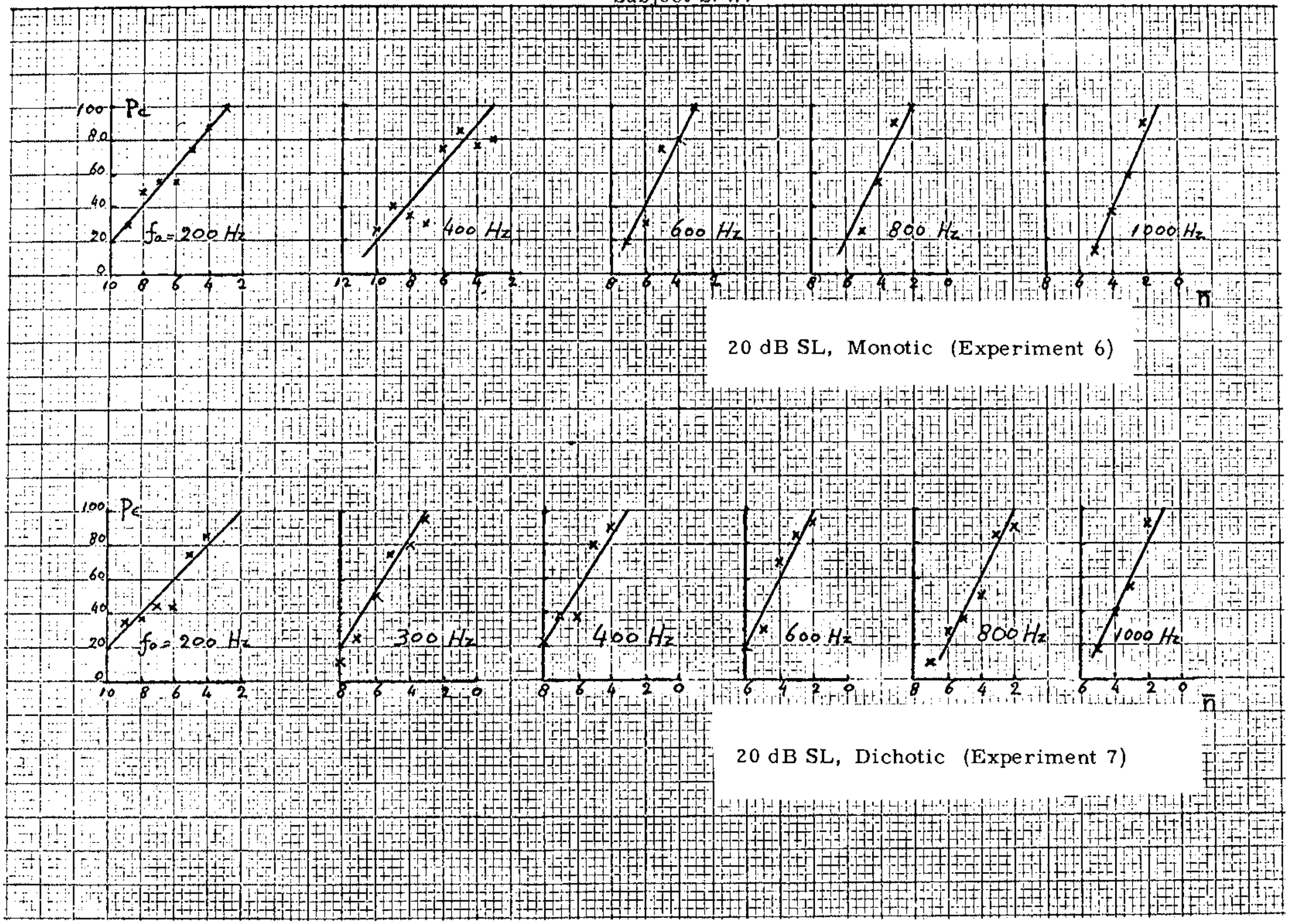


50 dB SPL, Dichotic (Experiment 9)



50 dB SPL, Dichotic, Simulated CT's (Experiment 10)

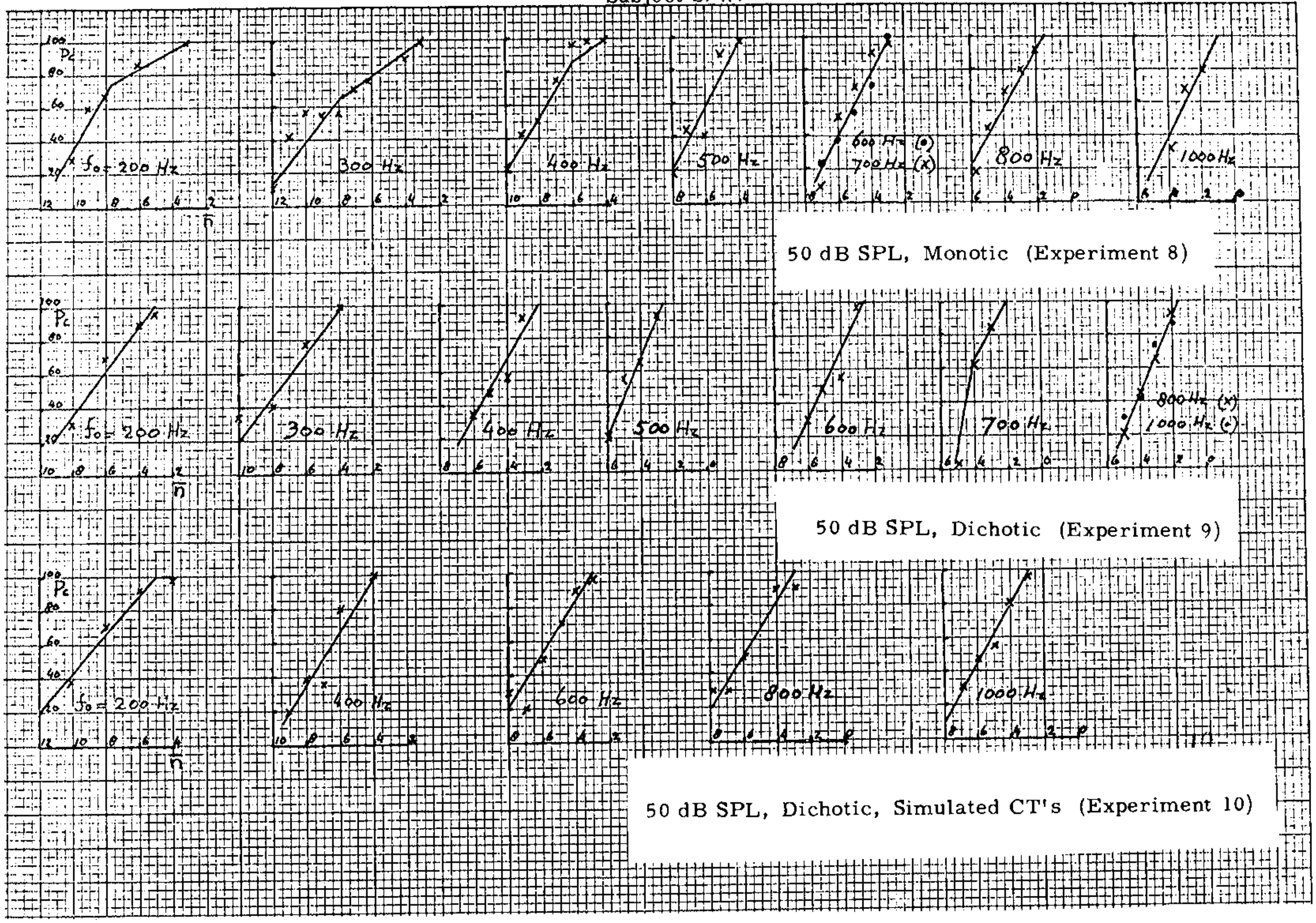
Subject S. W.



20 dB SL, Monotic (Experiment 6)

20 dB SL, Dichotic (Experiment 7)

Subject S. W.

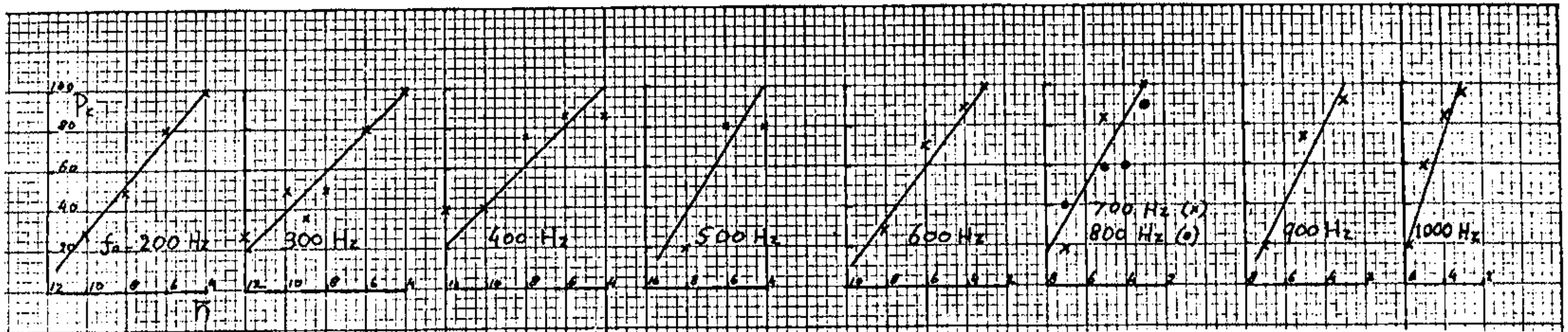


50 dB SPL, Monotic (Experiment 8)

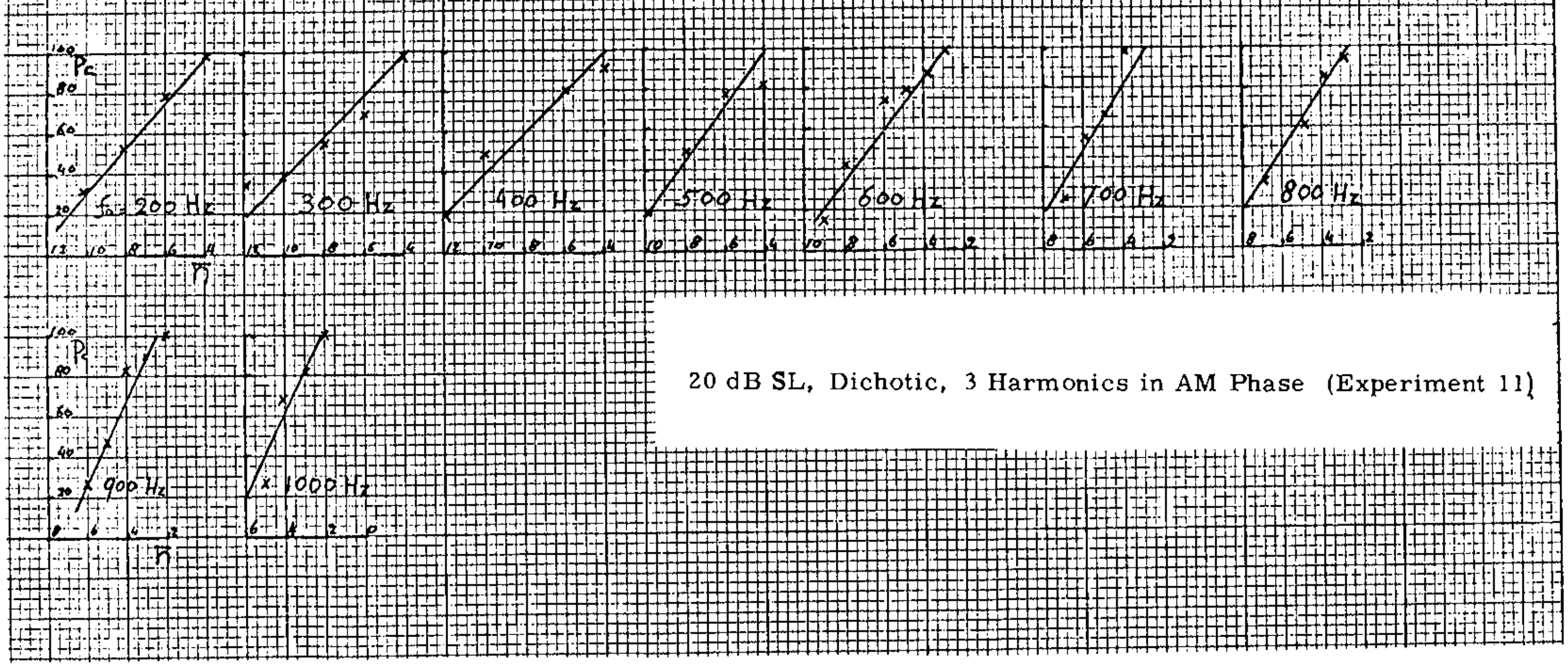
50 dB SPL, Dichotic (Experiment 9)

50 dB SPL, Dichotic, Simulated CT's (Experiment 10)

Subject S. M.

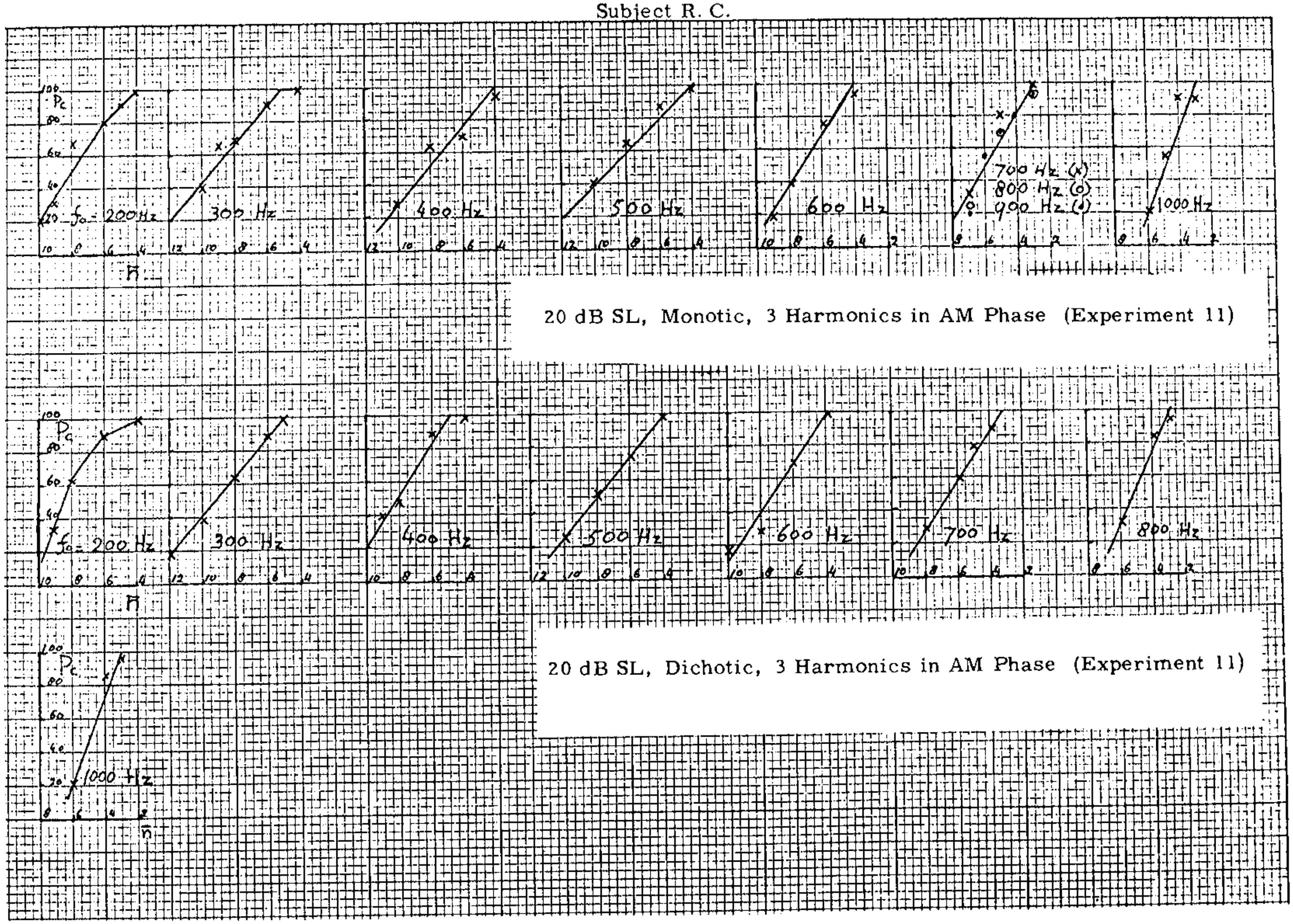


20 dB SL, Monotic, 3 Harmonics in AM Phase (Experiment 11)



20 dB SL, Dichotic, 3 Harmonics in AM Phase (Experiment 11)

Subject R. C.



20 dB SL, Monotic, 3 Harmonics in AM Phase (Experiment 11)

20 dB SL, Dichotic, 3 Harmonics in AM Phase (Experiment 11)

APPENDIX II

Assumptions for the Decision Model

A simple decision model can be described by the following axioms.

1. Each tone complex S_i is transformed into a unidimensional random decision variable P, having a mean $\overline{P}$, some small variance σ^2, and a unimodal density of arbitrary shape.

2. For M stimuli, there are M+1 criteria partitioning the decision space: $-\infty = C_0 \leqslant C_1 \leqslant \dots < C_M = \infty$.

3. Each response is determined by the particular partition in which P is enclosed in a given trial.

Applying this model to Experiment 17, we can make the following statements.

1. Only the second sound of each stimulus contains relevant information; hence, we can disregard the first sound.

2. Because of the random frequency shift, each stimulus category contains many possible complex tones, whose equivalent note or pitch values are given by the "first effect" relation $\Delta N = \Delta f/n$, where Δf is the amount of inharmonic frequency shift, and n is the harmonic number (which can be found by dividing a particular stimulus frequency by the difference frequency and then taking the closent integer). Since the amount of random frequency shift was limited to $\pm f_0/4$, where f_0 is the underlying fundamental of the first note, the stimulus input can be modeled as a random variable N, having a uniform distribution of width $f_0/2n$ and a mean determined by the fundamental of the second note.

If we assume that the sensation variance σ^2 is negligible compared with the variance of N, then the conditional probability density functions of P under the different stimulus conditions will be uniform, have equal width $(f_0/2n)$, and means that are approximately $f_0/16$ apart (semitone ratio), except for the distance between means for stimuli 4 and 5, which will be $f_0/8$ (full-tone ratio).

Finally, let us assume that the criteria C_1 through C_7 are placed in such a way that average percentage correct response is maximized; this means that they are located anywhere in the overlapping region of two adjacent density functions.

The average percentage correct responses, P_c, is then given by

$$P_c = \Pr\,[P < C_1/S_1]\,\Pr\,[S_1] + \Pr\,[C_1 \leqslant P < C_2/S_2]\,\Pr\,[S_2] + \dots + \Pr\,[P \geqslant C_7/S_8]\,\Pr\,[S_8].$$

This expression has to be evaluated in 3 different regions:

a. $n \geqslant 8$. The distance between the means for all density functions is larger than (or equal to) twice the standard deviation. There is no overlap, and hence the average score, P_c, equals 1.

b. $4 \leqslant n \leqslant 8$. All density functions overlap, except those for stimuli 4 and 5. The average score is given by

$$P_c = \frac{1}{4}\left(\frac{1}{2} + (f_0/32\times 2n/f_0)\right) + \frac{1}{2}(f_0/16\times 2n/f_0) + \frac{1}{4}\left(\frac{1}{2} + (f_0/32\times 2n/f_0)\right)$$

$$= (3n+8)/32.$$

c. $n \leqslant 4$. All densities overlap. The average score, P_c, is given by

$$P_c = \frac{1}{4}\left(\frac{1}{2} + (f_0/32\times 2n/f_0)\right) + \frac{1}{2}(f_0/16\times 2n/f_0) + \frac{1}{4}\left((f_0/32 + f_0/16)\times 2n/f_0\right)$$

$$= 1/8 + n/8.$$

Summarizing, we have

$$\begin{aligned} P_c &= 1 && \text{for } n \geqslant 8 \\ &= (3n+8)/32 && \text{for } 4 \leqslant n \leqslant 8 \\ &= 1/8 + n/8 && \text{for } n \leqslant 4. \end{aligned}$$

In this computation, we assumed merely for mathematical convenience that the criteria were placed symmetrically between the means. The same results will be obtained for the less restricted condition which was stated earlier. The piecewise linear relation between P_c and n is shown in Fig. 21, where n is replaced with $\bar{n}$, the average harmonic number randomly chosen over a range of three.

APPENDIX III

Definitions

The meanings of some musical terms in this report are somewhat different from standard definitions. To avoid confusion and to make the meaning of various terms clear, we have adopted the following definitions.

A note is a musical symbol, representing a unique fundamental frequency; it is equivalent to a number, and the time information, normally also contained in a musical note, is disregarded.

A melody is a sequence of notes. The term is used in a very wide sense and distinctions among melody, series or note sequence, often made by musicians, are disregarded.

A note scale is an ordered set of notes, having a defined one-to-one relationship with an ordered set of frequencies which has ratio properties.

A musical sound is the acoustical representation of a note; it is a periodic or quasi-periodic sound whose fundamental frequency equals that designated by the note.

A melody sensation is the subjective correlate of a sequence of musical sounds corresponding to notes on a musical scale.

Musical pitch is the subjective correlate of each of the musical sounds in a melody.

Acknowledgment

The authors wish to express their gratitude to many people who have contributed to this study. The laboratory facilities that made this work possible were developed by William F. Kelley, Dennis J. Callahan, and Louis D. Braida. Informative discussions were held with our colleagues H. Steven Colburn, Nathaniel I. Durlach, David S. Epstein, and William M. Siebert.

A special word of thanks goes to Susan Middleton, Ruth Chatman, Nicholas Humez, and Steven Weil, who patiently sat through hours of experiments and contributed their musical skill to the experimental part of this work.

Finally, we want to express our indebtedness to Jan F. Schouten, Director of the Institute for Perception Research, Eindhoven, The Netherlands, whose scientific work and leadership made us appreciate the importance of periodicity pitch – alias the "case of the missing fundamental."

References

1. M. Mersenne, Harmonie Universelle: The Books on Instruments. Translated into English by Roger E. Chapman (M. Nijhoff, The Hague, 1957).
2. Galileo Galilei, Dialogues Concerning Two New Sciences. Translated into English by Henry Crew and Alfonso de Salvio (The Macmillan Company, New York, 1914).
3. A. Seebeck, "Beobachtungen ueber einige Bedingungen der Entstehung von Tönen," Ann. Phys. Chem. 53, 417-436 (1841).
4. G. S. Ohm, "Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen," Ann. Phys. Chem. 59, 513-565 (1843).
5. G. S. Ohm, "Noch ein Paar Worte ueber die Definition des Tones," Ann. Phys. Chem. 62, 1-18 (1844).
6. A. Seebeck, "Ueber die Sirene," Ann. Phys. Chem. 60, 449-481 (1843).
7. A. Seebeck, "Ueber die Definition des Tones," Ann. Phys. Chem. 63, 353-368 (1844).
8. A. Seebeck, "Ueber die Erzeugung von Tonen durch getrennte Eindrücke, mit beziehung auf die Definition des Tones," Ann. Phys. Chem. 63, 368-380 (1844).
9. H. L. F. von Helmholtz, "Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik" (F. Vieweg und Sohn, Braunschweig, 1863). English Translation: "On the Sensations of Tone" (Dover Publications, Inc., New York, 1954).
10. G. Tartini, "Trattato di musica secondo la vera scienza dell'armonia" (Padua, 1754).
11. G. A. Sorge, "Anweisung zur Stimmung und Temperatur sowohl der Orgelwerke, als auch anderer Instrumente, sonderlich aber des Claviers" (Hamburg, 1744).
12. J. B. Romieu, "Nouvelle découverte des sons harmoniques graves dont la resonance est très sensible dans les accords des instruments à vent," Assemblée Publique de la Société Royale des Sciences, Montpelier, 1751.
13. J. W. Strutt, Baron Rayleigh, The Theory of Sound, Vol. II (Dover Publications, Inc., New York, 2d ed. (1896) revised and enlarged, 1945).
14. J. F. Schouten, "The Residue, a New Component in Subjective Sound Analysis," Proc. Kon. Nederl. Akad. Wetensch. 43, 356-365 (1940).
15. R. Plomp, "Experiments on Tone Perception," Doctoral Thesis, University of Utrecht, 1966.
16. R. L. Wegel and C. E. Lane, "The Auditory Masking of One Pure Tone by Another and Its Probable Relation to the Dynamics of the Inner Ear," Phys. Rev. 23, 266-285 (1924).
17. G. von Békésy, Experiments in Hearing (McGraw-Hill Book Company, Inc., New York, 1960).
18. R. Galambos and H. Davis, "The Response of Single Auditory-Nerve Fibres to Acoustic Stimulation," J. Neurophysiol. 6, 39-57 (1943).
19. R. Galambos and H. Davis, "Action Potentials from Single Auditory-Nerve Fibers," Science 108, 513 (1948).
20. I. Tasaki, "Nerve Impulses in Individual Auditory Nerve Fibers of Guinea Pig," J. Neurophysiol. 17, 97-122 (1954).
21. N. Y. S. Kiang, Discharge Patterns of Single Fibers in the Cat's Auditory Nerve, M.I.T. Research Monograph No. 35 (The M.I.T. Press, Cambridge, Mass., 1965).

22. G. von Békésy, "Ueber die Frequenzauflösung in der menschlichen Schnecke," Acta Oto-laryngol. 32, 60-84 (1944).

23. G. von Békésy, "Paradoxical Direction of Wave Travel along the Cochlear Partition," J. Acoust. Soc. Am. 27, 155-161 (1955).

24. E. Zwicker, "Der Ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohres," Acustica 5, 67-74 (1955).

25. J. L. Goldstein, "Auditory Spectral Filtering and Monaural Phase Perception," J. Acoust. Soc. Am. 41, 458-478 (1967).

26. J. L. Goldstein, "Auditory Nonlinearity," J. Acoust. Soc. Am. 41, 676-689 (1967).

27. J. L. Goldstein, "Aural Combination Tones," in R. Plomp and G. F. Smoorenburg (Eds.), Frequency Analysis and Periodicity Detection in Hearing (A. W. Sijthoff, Leiden, The Netherlands, 1970), pp. 230-247.

28. J. L. Goldstein, "Research Objectives: 4. Aural Combination Tones," Quarterly Progress Report No. 100, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., January 15, 1971, pp. 202-203.

29. J. L. Goldstein and N. Y. S. Kiang, "Neural Correlates of the Aural Combination Tone $2f_1-f_2$," Proc. IEEE 56, 981-992 (1968).

30. H. Fletcher, "The Physical Criterion for Determining the Pitch of a Musical Tone," Phys. Rev. 23, 427-437 (1924).

31. C. Stumpf, Tonpsychologie, Vol. 2 (Verlag S. Hirzel, Leipzig, 1890).

32. R. König, "Ueber den Zusammenklang zweier Töne," Ann. Phys. Chem. 157, 177-237 (1876). English translation: "On the Simultaneous Sounding of Two Notes," Phil. Mag., Ser. 5, Vol. I, pp. 417-446; 511-525, 1876.

33. L. Hermann, "Neue Versuche zur Frage der Unterbrechungstöne," Arch. ges. Physiol. 146, 249-294 (1912).

34. H. L. F. von Helmholtz, "Ueber den Muskelton," Verhandl. Naturhist. -Medic. Vereins, Heidelberg 4, 88 (1868).

35. J. Bernstein, "Untersuchungen ueber den Erregungsvorgang im Nerven- und Muskelsysteme" (Carl Winter's Universitätsbuchhandlung, Heidelberg, 1871).

36. S. P. Thompson, "On Binaural Audition," Phil. Mag., Ser. 5, Vol. 4, pp. 274-276, 1877.

37. H. W. Dove, "Repertorium der Physik," Vol. 3 (Verlag Veit & Comp., Berlin, 1839).

38. W. Wundt, "Grundzüge der Physiologischen Psychologie," Vol. I (Verlag Engelmann, Leipzig, 1880).

39. E. G. Wever and C. W. Bray, "Present Possibilities for Auditory Theory," Psychol. Rev. 37, 365-380 (1930).

40. W. Rutherford, "A New Theory of Hearing," J. Anat. Physiol. 21, 166-168 (1886).

41. M. Meyer, "Ueber Kombinationstöne und einige hierzu in Beziehung stehende akustische Erscheinungen," Z. Psychol. Physiol. Sinnesorgane 11, 177-229 (1896).

42. M. Meyer, "Zur Theorie der Differenztöne und der Gehörsempfindungen überhaupt," Z. Psychol. Physiol. Sinnesorgane 16, 1-34; 17, 8-14 (1898).

43. M. Meyer, "Zur Theorie des Hörens," Arch. ges. Physiol. 78, 346-362 (1899).

44. M. Meyer, "An Introduction to the Mechanics of the Inner Ear," University of Missouri Studies, Sci. Ser. 2, No. 1, 1907.

45. M. Meyer, "How We Hear" (Charles T. Branford Company, Newton, Mass., 1950).

46. J. F. Schouten, "The Perception of Subjective Tones," Proc. Kon. Nederl. Akad. Wetensch. 41, 1086-1093 (1938).

47. J. F. Schouten, "The Perception of Pitch," Philips Tech. Rev. 5, 286-294 (1940).

48. J. F. Schouten, "The Residue and the Mechanism of Hearing," Proc. Kon. Nederl. Akad. Wetensch. 43, 991-999 (1940).

49. E. de Boer, "On the Residue in Hearing," Doctoral Thesis, University of Amsterdam, 1956.

50. A. M. Small, Jr., "Some Parameters Influencing the Pitch of Amplitude Modulated Signals," J. Acoust. Soc. Am. 27, 751-760 (1955).

51. J. F. Schouten, R. J. Ritsma, and B. L. Cardozo, "Pitch of the Residue," J. Acoust. Soc. Am. 34, 1418-1424 (1962).

52. H. Fishler, "Model of the Secondary Residue Effect in the Perception of Complex Tones," J. Acoust. Soc. Am. 42, 759-764 (1967).

53. K. Walliser, "Zusammenwirkung von Hüllkurvenperiode und Tonheit bei der Bildung der Periodentonhöhe," Doctoral Thesis, Technische Hochschule, München, 1968.

54. M. R. Schroeder, "Residue Pitch: A Remaining Paradox and a Possible Explanation," J. Acoust. Soc. Am. 40, 79-81 (1966).

55. R. J. Ritsma, "Periodicity Detection," in R. Plomp and G. F. Smoorenburg (Eds.), Frequency Analysis and Periodicity Detection in Hearing (A. W. Sijthoff, Leiden, The Netherlands, 1970).

56. G. F. Smoorenburg, "Pitch Perception of Two-Frequency Stimuli," J. Acoust. Soc. Am. 48, 924-942 (1970).

57. R. J. Ritsma, "Existence Region of the Tonal Residue I.," J. Acoust. Soc. Am. 34, 1224-1229 (1962).

58. R. J. Ritsma, "Existence Region of the Tonal Residue II.," J. Acoust. Soc. Am. 35, 1241-1245 (1963).

59. R. J. Ritsma, "Frequencies Dominant in the Perception of Complex Sounds," J. Acoust. Soc. Am. 42, 191-198 (1967).

60. R. J. Ritsma, "Frequencies Dominant in Pitch of Periodic Pulses of Alternating Polarity," IPO Ann. Progr. Rep. 2, 14-24, Eindhoven (1967).

61. W. R. Thurlow, "Perception of Low Auditory Pitch: A Multicue Mediation Theory," Psych. Rev. 70, 461-470 (1963).

62. (a) W. D. Ward, "Subjective Musical Pitch," J. Acoust. Soc. Am. 26, 369-380 (1954).

    (b) A. J. M. Houtsma, "Auditory Discrimination of Frequency Ratios," J. Acoust. Soc. Am. 44, 383 (1968).

63. M. B. Sachs and N. Y. S. Kiang, "Two-Tone Inhibition in Auditory-Nerve Fibers," J. Acoust. Soc. Am. 43, 1120-1128 (1968).

64. M. B. Sachs, "Stimulus Response Relation for Auditory-Nerve Fibers: Two-Tone Stimuli," J. Acoust. Soc. Am. 45, 1025-1036 (1969).

65. R. J. Ritsma, B. L. Cardozo, and A. Cohen, "Acoustic Demonstration to Perception of Pitch; Demonstration to Phonetics," Phonograph record in Philips Techn. Tijdschrift, Vol. 24, p. 341, 1962.

66. D. Cross and H. Lane, "Attention to Single Stimulus Properties in the Identification of Complex Tones," in Experimental Analysis of the Control of Speech Production and Perception, ORA Report No. 05613-1-P, University of Michigan, Ann Arbor, Mich., 1963.

67. S. P. Thompson, "Phenomena of Binaural Audition, Part II," Phil. Mag., Ser. 5, Vol. 6, pp. 383-391, 1878.

68. S. P. Thompson, "Phenomena of Binaural Audition, Part III," Phil. Mag., Ser. 5, Vol. 12, pp. 351-355, 1881.

69. W. R. Thurlow and A. M. Small, Jr., "Pitch Perception for Certain Periodic Auditory Stimuli," J. Acoust. Soc. Am. 27, 132-137 (1955).

70. J. L. Flanagan and N. Guttman, "On the Pitch of Periodic Pulses," J. Acoust. Soc. Am. 32, 1308-1319 (1960).

71. J. L. Flanagan and N. Guttman, "Pitch of Periodic Pulses without Fundamental Components," J. Acoust. Soc. Am. 32, 1319-1328 (1960).

72. R. C. Mathes and R. L. Miller, "Phase Effects in Monaural Perception," J. Acoust. Soc. Am. 19, 780-797 (1947).

73. J. C. R. Licklider, "Auditory Frequency Analysis," Third London Symposium on Information Theory, 1955.

74. G. A. Hoogland, "The Missing Fundamental," Doctoral Thesis, University of Utrecht, 1953.

75. R. F. Ritsma and F. L. Engel, "Pitch of Frequency-Modulated Signals," J. Acoust. Soc. Am. 36, 1637-1644 (1964).

76. W. Thomson, Baron Kelvin, "On Beats of Imperfect Harmonies," Proc. Roy. Soc. Edinburgh 9, 602-612 (1878).

77. G. A. Miller, "The Magical Number Seven Plus or Minus Two," Psych. Rev. 63, 81-97 (1956).

78. L. L. Thurstone, "A Law of Comparative Judgement," Psych. Rev. 34, 273-286 (1927).

79. L. L. Thurstone, "Three Psychophysical Laws," Psych. Rev. 34, 424-432 (1927).

80. N. I. Durlach and L. D. Braida, "Intensity Perception I. Preliminary Theory of Intensity Resolution," J. Acoust. Soc. Am. 46, 372-383 (1969).

81. E. G. Shower and R. Biddulph, "Differential Pitch Sensitivity of the Ear," J. Acoust. Soc. Am. 3, 275-287 (1931).

82. T. Koester, "The Time Error and Sensitivity in Pitch and Loudness Discrimination as a Function of Time Interval and Stimulus Level," Arch. Psychol. N.Y., Vol. 41, No. 297, pp. 1-69, 1945.

83. J. D. Harris, "Pitch Discrimination," J. Acoust. Soc. Am. 24, 750-755 (1952).

84. A. Bachem, "Various Types of Absolute Pitch," J. Acoust. Soc. Am. 9, 146-151 (1937).

85. P. T. Brady, "Fixed-Scale Mechanism of Absolute Pitch," J. Acoust. Soc. Am. 48, 883-887 (1970).

86. E. Mach, "Zur Analyse der Tonempfindungen," Sitzungsber. Kaisersl. Akad. Wissensch., Mathemat.-Naturwissensch., Klasse 92, II Abt., 1881. English translation: "The Analysis of Sensations" (Dover Publications, Inc., New York, 1959).

87. W. D. Ward, "Absolute Pitch, Parts I and II," Sound, Vol. 2, No. 3, pp. 14-21; Vol. 2, No. 4, pp. 33-41, 1963.

88. H. L. F. von Helmholtz, *On the Sensations of Tone*, *op. cit.*, Foreword to 3rd German edition.

89. E. Terhardt, "Frequency Analysis and Periodicity Detection in the Sensations of Roughness and Periodicity Pitch," in R. Plomp and G. F. Smoorenburg (Eds.), *Frequency Analysis and Periodicity Detection in Hearing* (A. W. Sijthoff, Leiden, The Netherlands, 1970).

90. W. M. Siebert, "Some Implications of the Stochastic Behavior of Primary Auditory Neurons," Kybernetik 2, 206-215 (1965).

91. P. R. Gray, "A Statistical Analysis of Electrophysiological Data from Auditory Nerve Fibers in Cat," Technical Report 451, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., June 21, 1966.

92. J. E. Evans, Private communication.

93. H. S. Colburn, "Some Physiological Limitations on Binaural Performance," Ph.D. Thesis, Massachusetts Institute of Technology, 1969.

94. J. E. Rose, J. F. Brugge, D. J. Anderson, and J. E. Hind, "Some Possible Neural Correlates of Combination Tones," J. Neurophysiol. 32, 402-423 (1969).

JOINT SERVICES ELECTRONICS PROGRAM
REPORTS DISTRIBUTION LIST

Department of Defense

Assistant Director (Research)
Office of Director of Defense, Research
& Engineering
The Pentagon, Rm 3C128
Washington, D.C. 20301

Technical Library
DDR&E
Room 3C-122, The Pentagon
Washington, D.C. 20301

Director For Materials Sciences
Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

Chief, R&D Division (340)
Defense Communications Agency
Washington, D.C. 20305

Defense Documentation Center
Attn: DDC-TCA
Cameron Station
Alexandria, Virginia 22314

Dr. Alvin D. Schnitzler
Institute For Defense Analyses
Science and Technology Division
400 Army-Navy Drive
Arlington, Virginia 22202

LTC Norman D. Jorstad
Weapons Systems Evaluation Group
400 Army-Navy Drive
Arlington, Virginia 22202

Central Intelligence Agency
Attn: CRS/ADD/PUBLICATIONS
Washington, D.C. 20505

Department of the Air Force

Hq USAF (AF/RDPE)
Washington, D.C. 20330

Hq USAF/RDPS
Washington, D.C. 20330

Hq USAF (RDSD)
The Pentagon
Washington, D.C. 20330

Colonel E. P. Gaines, Jr.
ESD (MCD)
L. G. Hanscom Field
Bedford, Massachusetts 01730

Dr L. A. Wood, Director
Electronic and Solid State Sciences
Air Force Office of Scientific Research
1400 Wilson Boulevard
Arlington, Virginia 22209

Dr Harvey E. Savely, Director
Life Sciences
Air Force Office of Scientific Research
1400 Wilson Boulevard
Arlington, Virginia 22209

AFSC (CCJ/Mr I. R. Mirman)
Andrews Air Force Base,
Washington, D.C. 20331

Rome Air Development Center
Attn: Documents Library (TDLD)
Griffiss Air Force Base, New York 13440

Mr H. E. Webb, Jr. (ISCP)
Rome Air Development Center
Griffiss Air Force Base, New York 13440

Dr L. M. Hollingsworth
AFCRL (CA)
L. G. Hanscom Field
Bedford, Massachusetts 01730

Hq ESD (TRI)
L. G. Hanscom Field
Bedford, Massachusetts 01730

Professor R. E. Fontana, Head
Dept of Electrical Engineering
Air Force Institute of Technology
Wright-Patterson Air Force Base,
Ohio 45433

Dr H. V. Noble, AFAL/TE
Chief, Electronics Technology Division
Air Force Avionics Laboratory
Wright-Patterson Air Force Base,
Ohio 45433

Director
Air Force Avionics Laboratory
Wright-Patterson Air Force Base,
Ohio 45433

AFAL/TEA (Mr R. D. Larson)
Wright-Patterson Air Force Base,
Ohio 45433

Director of Faculty Research
Department of the Air Force
U.S. Air Force Academy
Colorado 80840

Mr Jules I. Wittebort
Chief, Electronics Branch
Manufacturing Technology Division
AFAL/LTE
Wright-Patterson Air Force Base,
Ohio 45433

Academy Library (DFSLB)
USAF Academy, Colorado 80840

Director of Aerospace Mechanics Sciences
Frank J. Seiler Research
Laboratory (OAR)
USAF Academy, Colorado 80840

Major Richard J. Gowen
Tenure Associate Professor
Dept of Electrical Engineering
USAF Academy, Colorado 80840

Director, USAF PROJECT RAND
Via: Air Force Liaison Office
The RAND Corporation
Attn: Library D
1700 Main Street
Santa Monica, California 90406

AUL3T-9663
Maxwell Air Force Base, Alabama 36112

AFETR Technical Library
(MU-135)
Patrick Air Force Base, Florida 32925

ADTC (SSLT)
Eglin Air Force Base, Florida 32542

Hq AMD (RDR/Lt Col Godden)
Brooks Air Force Base, Texas 78235

USAFSAM (RAT)
Brooks Air Force Base, Texas 78235

Commanding General
Attn: STEWS-AD-L, Technical Library
White Sands Missile Range,
New Mexico 88002

European Office of Aerospace Research
Technical Information Office
Box 14, FPO New York 09510

VELA Seismological Center
312 Montgomery Street
Alexandria, Virginia 22314

Capt C. E. Baum
AFWL (SRE)
Kirtland Air Force Base,
New Mexico 87117

Biomedical Engineering Branch (SCB)
USAF School of Aerospace Medicine (AFSC)
Department of the Air Force
Brooks Air Force Base, Texas 78235

Department of the Army

Director
Physical & Engineering Sciences Division
3045 Columbia Pike
Arlington, Virginia 22204

Commanding General
U.S. Army Security Agency
Attn: IARD-T
Arlington Hall Station
Arlington, Virginia 22212

Commanding General
U.S. Army Materiel Command
Attn: AMCRD-TP (Dr Zarwyn)
Washington, D.C. 20315

Director
U.S. Army Advanced Materiel
Concepts Agency
2461 Eisenhower Avenue
Alexandria, Virginia 22314

JOINT SERVICES REPORTS DISTRIBUTION LIST (continued)

Commanding General
USACDC Institute of Land Combat
Attn: Technical Library, Rm 636
2461 Eisenhower Avenue
Alexandria, Virginia 22314

Mr H. T. Darracott (AMXAM-FT)
U.S. Army Advanced Materiel
Concepts Agency
2461 Eisenhower Avenue
Alexandria, Virginia 22314

Dr Berthold Altmann (AMXDO-TI)
Harry Diamond Laboratories
Connecticut Avenue and
Van Ness Street N. W.
Washington, D.C. 20438

Commanding Officer (AMXRD-BAD)
U.S. Army Ballistic Research Laboratory
Aberdeen Proving Ground
Aberdeen, Maryland 21005

U.S. Army Munitions Command
Attn: Science & Technology Information
Branch, Bldg 59
Picatinny Arsenal, SMUPA-RT-S
Dover, New Jersey 07801

Dr Herman Robl
Deputy Chief Scientist
U.S. Army Research Office (Durham)
Box CM, Duke Station
Durham, North Carolina 27706

Richard O. Ulsh (CRDARD-IP)
U.S. Army Research Office (Durham)
Box CM, Duke Station
Durham, North Carolina 27706

Technical Director (SMUFA-A2000-107-1)
Frankford Arsenal
Philadelphia, Pennsylvania 19137

Redstone Scientific Information Center
Attn: Chief, Document Section
U.S. Army Missile Command
Redstone Arsenal, Alabama 35809

Commanding General
U.S. Army Missile Command
Attn: AMSMI-RR
Redstone Arsenal, Alabama 35809

Commanding General
U.S. Army Strategic Communications
Command
Attn: SCC-ATS (Mr Peter B. Pichetto)
Fort Huachuca, Arizona 85613

Dr Homer F. Priest
Chief, Materials Sciences Division,
Bldg 292
Army Materials & Mechanics Research
Center
Watertown, Massachusetts 02172

Commandant
U.S. Army Air Defense School
Attn: Missile Science Division, C&S Dept
P. O. Box 9390
Fort Bliss, Texas 79916

Commandant
U.S. Army Command and General
Staff College
Attn: Acquisitions, Lib Div
Fort Leavenworth, Kansas 66027

Dr Hans K. Ziegler
Army Member TAC/JSEP (AMSEL-XL-D)
U.S. Army Electronics Command
Fort Monmouth, New Jersey 07703

Mr I. A. Balton, AMSEL-XL-D
Executive Secretary, TAC/JSEP
U.S. Army Electronics Command
Fort Monmouth, New Jersey 07703

Director (NV-D)
Night Vision Laboratory, USAECOM
Fort Belvoir, Virginia 22060

Commanding Officer
Atmospheric Sciences Laboratory
U.S. Army Electronics Command
White Sands Missile Range,
New Mexico 88002

Atmospheric Sciences Laboratory
U.S. Army Electronics Command
Attn: AMSEL-BL-DD (Mr Marvin Diamond)
White Sands Missile Range,
New Mexico 88002

Chief
Missile Electronic Warfare Tech
Area (AMSEL-WL-M)
Electronic Warfare Laboratory,
USAECOM
White Sands Missile Range,
New Mexico 88002

Project Manager NAVCON
Attn: AMCPM-NC, Bldg 439
(H. H. Bahr)
Fort Monmouth, New Jersey 07703

Mr A. D. Bedrosian, Rm 26-131
U.S. Army Scientific Liaison Office
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, Massachusetts 02139

Commanding General
U.S. Army Electronics Command
Fort Monmouth, New Jersey 07703
Attn: AMSEL-RD-O (Dr W. S. McAfee)
GG-DD
XL-D
XL-DT
XL-G (Dr S. Kronenberg)
XL-H (Dr R. G. Buser)
BL-FM-P
CT-D
CT-R
CT-S
CT-L
CT-O
CT-I
CT-A
NL-D (Dr H. Bennett)
NL-A
NL-C
NL-P
NL-P-2
NL-R
NL-S
KL-D
KL-I
KL-E
KL-S
KL-SM
KL-T
VL-D
VL-F
WL-D

Department of the Navy

Director, Electronics Programs
Attn: Code 427
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Mr Gordon D. Goldstein, Code 437
Information Systems Program
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Commander
Naval Security Group Command
Naval Security Group Headquarters
Attn: Technical Library (G43)
3801 Nebraska Avenue, N. W.
Washington, D.C. 20390

Director
Naval Research Laboratory
Washington, D.C. 20390
Attn: Code 5200
Mr A. Brodzinsky, Supt,
Electronics Div

Director
Naval Research Laboratory
Attn: Dr H. Rabin, Code 7000
Washington, D.C. 20390

Code 8050
Maury Center Library
Naval Research Laboratory
Washington, D.C. 20390

Dr G. M. R. Winkler
Director, Time Service Division
U.S. Naval Observatory
Washington, D.C. 20390

Naval Air Systems Command
AIR 310 - Research Administrator
Room 424 JP-1
Washington, D.C. 20360

Dr A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps (Code AX)
Washington, D.C. 20380

Naval Ship Systems Command
Ship 035
Washington, D.C. 20360

U.S. Naval Weapons Laboratory
Dahlgren, Virginia 22448

JOINT SERVICES REPORTS DISTRIBUTION LIST (continued)

Naval Electronic Systems Command
Attn: Code 0311, Rm 7W12, NC #1
Department of the Navy
Washington, D.C. 20360

Commander
U.S. Naval Ordnance Laboratory
Attn: Librarian
White Oak, Maryland 20910

Director
Office of Naval Research
Boston Branch
495 Summer Street
Boston, Massachusetts 02210

Commander (ADL)
Naval Air Development Center
Attn: NADC Library
Johnsville, Warminster,
Pennsylvania,18974

Commanding Officer
Naval Missile Center
Attn: 5632.2, Technical Library
Point Mugu, California 93042

W. A. Eberspacher, Associate Head
Systems Integration Division, Code 5340A
U.S. Naval Missile Center
Point Mugu, California 93041

Commander
Naval Electronics Laboratory Center
Attn: Library
San Diego, California 92152

Deputy Director and Chief Scientist
Office of Naval Research Branch Office
1031 East Green Street
Pasadena, California 91101

Library (Code 2124)
Naval Postgraduate School
Monterey, California93940

Officer in Charge, New London Lab.
Naval Underwater Systems Center
Attn: Technical Library
New London, Connecticut 06320

Commanding Officer
Naval Avionics Facility
Attn: D/035 Technical Library
Indianapolis, Indiana 46241

Director
Naval Research Laboratory
Attn: Library, Code 2629 (ONRL)
Washington, D.C. 20390

Commanding Officer
Naval Training Device Center
Attn: Technical Library
Orlando, Florida 32813

U.S. Naval Oceanographic Office
Attn: M. Rogofsky, Librarian (Code 1640)
Washington, D.C. 20390

Other Government Agencies

Dr H. Harrison
AEC/NASA Space Nuclear Systems
Office
AEC Headquarters (F-309)
Washington, D.C. 20545

NASA Lewis Research Center
Attn: Library
21000 Brookpark Road
Cleveland, Ohio 44135

Los Alamos Scientific Laboratory
Attn: Reports Library
P. O. Box 1663
Los Alamos, New Mexico 87544

Mr M. Zane Thornton
Deputy Director
Center for Computer Sciences and
Technology
National Bureau of Standards
U.S. Department of Commerce
Washington D.C. 20234

U.S. Postal Service
Library – Room 6012
12th & Pennsylvania Ave., N. W.
Washington, D.C. 20260

Non-Government Agencies

Director
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Mr Jerome Fox, Research Coordinator
Polytechnic Institute of Brooklyn
333 Jay Street
Brooklyn, New York 11201

Director
Columbia Radiation Laboratory
Columbia University
538 West 120th Street
New York, New York 10027

Director
Coordinate Science Laboratory
University of Illinois
Urbana, Illinois 61801

Director
Stanford Electronics Laboratory
Stanford University
Stanford, California 94305

Director
Microwave Laboratory
Stanford University
Stanford, California 94305

Director
Electronics Research Laboratory
University of California
Berkeley, California 94720

Director
Electronics Sciences Laboratory
University of Southern California
Los Angeles, California 90007

Director
Electronics Research Center
The University of Texas at Austin
Engineering-Science Bldg 110
Austin, Texas 78712

Division of Engineering and
Applied Physics
210 Pierce Hall
Harvard University
Cambridge, Massachusetts 02138

Dr G. J. Murphy
The Technological Institute
Northwestern University
Evanston, Illinois 60201

Dr John C. Hancock, Head
School of Electrical Engineering
Purdue University
Lafayette, Indiana 47907

Dept of Electrical Engineering
Texas Technological University
Lubbock, Texas 79409

Aerospace Corporation
P. O. Box 95085
Attn: Library Acquisitions Group
Los Angeles, California 90045

Airborne Instruments Laboratory
Deerpark, New York 11729

The University of Arizona
Department of Electrical Engineering
Tucson, Arizona 85721

Chairman, Electrical Engineering
Arizona State University
Tempe, Arizona 85281

Engineering and Mathematical
Sciences Library
University of California at Los Angeles
405 Hilgred Avenue
Los Angeles, California 90024

Sciences-Engineering Library
University of California
Santa Barbara, California 93106

Professor Nicholas George
California Institute of Technology
Pasadena, California 91109

Aeronautics Library
Graduate Aeronautical Laboratories
California Institute of Technology
1201 E. California Boulevard
Pasadena, California 91109

Hunt Library
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213

Dr A. G. Jordan
Head of Dept of Electrical Engineering
Carnegie-Mellon University
Pittsburg, Pennsylvania 15213

Case Western Reserve University
Engineering Division
University Circle
Cleveland, Ohio 44106

Hollander Associates
Attn: Librarian
P. O. Box 2276
Fullerton, California 92633

Walter H. Veazie, Head
Electronic Properties Information Center
Mail Station E-148
Hughes Aircraft Company
Culver City, California 90230

Illinois Institute of Technology
Department of Electrical Engineering
Chicago, Illinois 60616

Government Documents Department
University of Iowa Libraries
Iowa City, Iowa 52240

The Johns Hopkins University
Applied Physics Laboratory
Attn: Document Librarian
8621 Georgia Avenue
Silver Spring, Maryland 20910

Lehigh University
Department of Electrical Engineering
Bethlehem, Pennsylvania 18015

Mr E. K. Peterson
Lenkurt Electric Co. Inc.
1105 Country Road
San Carlos, California 94070

MIT Lincoln Laboratory
Attn: Library A-082
P. O. Box 73
Lexington, Massachusetts 02173

Miss R. Joyce Harman
Project MAC, Room 810
545 Main Street
Cambridge, Massachusetts 02139

Professor R. H. Rediker
Electrical Engineering, Professor
Massachusetts Institute of Technology
Building 13-3050
Cambridge, Massachusetts 02139

Professor Joseph E. Rowe
Chairman, Dept of Electrical Engineering
The University of Michigan
Ann Arbor, Michigan 48104

New York University
Engineering Library
Bronx, New York 10453

Professor James A. Cadzow
Department of Electrical Engineering
State University of New York at Buffalo
Buffalo, New York 14214

Department of Electrical Engineering
Clippinger Laboratory
Ohio University
Athens, Ohio 45701

Raytheon Company
Research Division Library
28 Seyon Street
Waltham, Massachusetts 02154

Rice University
Department of Electrical Engineering
Houston, Texas 77001

Dr Leo Young, Program Manager
Stanford Research Institute
Menlo Park, California 94025

GTE Waltham Research Lab Library
Attn: Documents Librarian
Esther McLaughlin
40 Sylvan Road
Waltham, Massachusetts 02154

Dr W. R. LePage, Chairman
Department of Electrical Engineering
Syracuse University
Syracuse, New York 13210

Dr F. R. Charvat
Union Carbide Corporation
Materials Systems Division
Crystal Products Department
8888 Balboa Avenue
P. O. Box 23017
San Diego, California 92123

Utah State University
Department of Electrical Engineering
Logan, Utah 84321

Research Laboratories for the
Engineering Sciences
School of Engineering and Applied Science
University of Virginia
Charlottesville, Virginia 22903

Yale University
Engineering and Applied Science Library
15 Prospect Street
New Haven, Connecticut 06520

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Research Laboratory of Electronics Massachusetts Institute of Technology Cambridge, Massachusetts 02139	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP None

3. REPORT TITLE
Perception of Musical Intervals: Evidence for the Central Origin of the Pitch of Complex Tones

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Technical Report

5. AUTHOR(S) *(First name, middle initial, last name)*
Adrianus J. M. Houtsma

Julius L. Goldstein

6. REPORT DATE October 1, 1971	7a. TOTAL NO. OF PAGES 90	7b. NO. OF REFS 94

8a. CONTRACT OR GRANT NO. DAABO7-71-C-0300 b. PROJECT NO. C8-1-01504-01-C8-CA c. NIH Grant 5 PO1 GM14940-05 d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report 484 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* None

10. DISTRIBUTION STATEMENT
THIS DOCUMENT HAS BEEN APPROVED FOR PUBLIC RELEASE AND SALE; ITS DISTRIBUTION IS UNLIMITED.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Joint Services Electronics Program Through U.S. Army Electronics Command

13. ABSTRACT

Melodies can be recognized in music, regardless of the instrument on which they are played. This is true even when the musical sounds have no acoustical energy at the fundamental frequency. This phenomenon was investigated qualitatively and quantitatively through a series of experiments in which subjects were asked to identify melodies and simple musical intervals. Each musical note was played by a complex tone comprising successive upper harmonics with randomly chosen lower harmonic number. Melodies and intervals played with such sounds consisting of only two partials of low harmonic number could be identified perfectly both when the complex tones were presented monotically (both partials to one ear) and when the partials were distributed dichotically (one partial to each ear). Control experiments showed that neither difference tone nor transformations based on frequency difference per se explain these phenomena. Percent correct identification decreased both with increasing fundamental frequency and average harmonic number. Performance is essentially the same for monotic and dichotic stimulus paradigms, except for differences which were shown to be accounted for by aural combination tones. Moreover, identification performance is essentially random when the harmonic numbers of the stimulus tones or audible aural combination tones are sufficiently high (greater than 10) so that they cannot be resolved behaviorally in monaural experiments. These findings suggest that sensations of "musical pitch of the fundamental" in complex tones are mediated centrally by neural signals derived from those stimulus partials that are tonotopically resolved, rather than being mediated by neural transformations of those upper partials which the peripheral auditory system fails to resolve.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
auditory frequency resolution						
aural combination tones						
aural phase effects						
binaural auditory mechanisms						
central auditory mechanisms						
musical consonance						
musical pitch						
music perception						
periodicity pitch						
pitch of complex tones						
pitch of missing fundamental						
Psychoacoustics						
Psychological and Physiological Acoustics						