Medical Research Archives

European Society of Medicine

OPEN ACCESS

Published: December 31, 2023

Citation: Essien, A. J., 2023. Pitch and Tone: Primacy of sound sources in auditory perception. Medical Research Archives, [online] 11(12). https://doi.org/10.18103/mra.v11i12.4828

DOI: https://doi.org/10.18103/mra.v11i12.4828

ISSN: 2375-1924

RESEARCH ARTICLE

Pitch and Tone: Primacy of sound sources in auditory perception

Akpan Jimmy Essien[1,2]

[1]Ecological Auditory Research Laboratory. London, United Kingdom
[2]Acoustical Society of Nigeria, University of Nigeria, Nsukka, Nigeria.

jessien@earlab.ac

ABSTRACT

The brain performs diverse and complex operations effortlessly. Its sensitivity to minute stimulus changes, and precision in scaling sensory magnitudes, are yet to be understood by scientists. Nevertheless, its vast vault is accessible through well-established psycho-physical principles for dealing with sensory, mental, and all intangible phenomena. However, hearing research is experiencing the scarcity of law-abiding scientists because the current generation of information-processing psychologists cannot discipline itself to abide by the laws that govern hearing research. Modern investigators hope to resolve all problems in auditory perception through random and effortless computations of endless varieties of meaningless variables without the concept of invariance. Consequently, pitch and tone in music or speech are still shrouded in mystery after 2,500 years of intensive inquiry. The failure augurs the existence of fundamental scientific errors. This paper describes a search for the origin of pitch which plays lexical and intonational roles in Yoruba, a tone language. The approach which is based on the ecological conception of sound, will furnish experimental evidence from speech surrogates (drums) and strings, guided by the concept of invariance. The investigations are designed to demonstrate that the brain does not perform any computations at all in pitch and tone perception. The results show that the brain measures a unique mechanical parameter of sound sources to which all physical and mechanical transformations lead. In every pitch/tone perception context, therefore, the brain discards all other factors which convey other attributes of sound outside pitch and tone. The finding accounts for constancy in auditory code perception despite overwhelming variability at the production and acoustic levels. The experimental data accord primacy to the sound source as the origin of auditory codes. Implications for future research at the mechanical, acoustic, and neurophysiological levels of investigations into pitch/tone and auditory analysis as a whole are discussed.

1. Introduction

Scientific investigations in music, speech, and hearing began at the mechanical level. Pythagoras (6th century B.C.) discovered a functional relationship between string ratios and musical pitch intervals. The discovery constitutes the hub around which turn all work in hearing[1]. For another 2,000 years, the theory gave rise to speculations about the universe, and sound was involved in "a semi-mystical arithmetic of music[2(p1)]." Thereafter, modern psychoacoustic principles in hearing research, based on Ohm's acoustic law[3] and Helmholtz's resonance/place theory of hearing [4] came on the scene, and frequency ratios replaced string ratios. Upon the new acoustic platform, Alexander Graham Bell built his life-long ambition to integrate the deaf and the hard-of-hearing into the hearing world. His efforts yielded the telephone. Today, in just under 150 years after, the telephone links every human on this earth and even with outer space. However, the same acoustic principle that expedited success in telephone engineering has failed to explain the pitch of the sound generated by an acoustic system as simple as a stretched string[5]. Why? That is the defiant and long-standing question that this paper addresses and aims to resolve. To that end, it is necessary to understand the problem.

Acoustics was conceived to explain auditory perception. Its success in other peripheral fields cannot compensate for its total failure to fulfil its original purpose[6,7]. The failure did not take scientists by surprise because it was "always clear that the ear does not react like a simple Fourier frequency analyzer[8(p631)]." The failure is a clear indication that principles of sound transmission do not apply in neurophysiological processes in auditory perception. The problem of non-invariance between frequency and pitch questions the optimism of scientists who hope not only to explain pitch but also speech, arguably the most complex of human behaviours, by information-processing procedures. Thus, the error lies wholly with modern hearing scientists who know (or should know) that any hearing theory that is not based on invariance has already failed before any efforts at putting it into practice[9,10]. Cognizant of this fact, many conferenciers have underscored the significance of invariance in speech and hearing research. Yet, there is not one single known work in music, speech, or auditory perception that incorporates invariance in its theory or practices. The result is failure because hearing scientists consciously (or otherwise) refuse to abide by the laws of their science.

Sound conveys several auditory attributes of which pitch is the primary and often the dominant one. The stochastic nature of sound suggests that the study of each auditory quality, for example, pitch, cannot succeed without invariance. The investigations reported herein, guided by invariance, will make no sense without acquaintance with this pivotal psycho-physical pre-requisite for dealing scientifically with sensation. Thus, this paper will briefly introduce the concept of invariance as it applies in auditory psychophysics[11,12]. Thereafter, the search for the invariant in tone (speech) and pitch (music) will be presented. Because this study addresses the pitch of natural sounds via sound production criteria, the stimuli comprise naturally-produced signals on drums and musical strings. The goal is two-fold: (1) Demonstrate the reason for the failure of psychoacoustics in hearing research, (2) Propose a new course to follow founded on invariance. Implications for future research in acoustics, neurophysiology, and auditory perception as a whole are discussed.

2. MATERIALS AND METHODS

2.1. INVARIANCE IN VARIABILITY

Invariance and its significance in speech, music acoustics, and perception has featured in many publications[10,12,13]. However, discussions at such conferences have never presented a clear understanding of invariance nor its application to guide hearing research. This is least surprising because although our entire life is driven by invariance, the thought of something that does not change when everything in the environment is changing is very challenging[13,14]. To establish a basis for this study, it is necessary to show that all perception theories of music, speech, and hearing have failed owing to the absence of invariance. First, we will examine how *invariance* guides us in everything we do, the significance of a physical correlate, and the distinction between a physical correlate and the stimulus to a sensation. These topics that establish the basis for this paper are discussed and demonstrated in more detail elsewhere[11].

Consider figure 1. It presents a sheet of paper cut in two portions. Portion A weighs 65g, and portion B weighs 35g. We conclude from relative judgment that the larger portion A weighs more than the smaller portion B. True. For invariance, however, the facts on hand are insufficient to conclude that A weighs more than B because invariance cannot be determined through simple visual observations and relative judgments. Invariance rests on a commonality of the observed entities which may not be open to visual inspection. Suppose, then, that we cut portion A1 (10 mm^2) and portion B1 (10 mm^2), and both portions A1 and B1 have the same weight (for example 15g). For the concept of invariance, the larger portion A weighing 65g, and the smaller portion B weighing 35g, have precisely the *same* weight. The established constant provides a basis for dealing scientifically with variability. Thus, we may enter any stationery store anywhere in the world and ask for paper, e.g., 80g sq/m, we are certain to obtain the same paper quality in any size, colour, new or recycled paper, matt or glossy, lined or plain, etc. Amid the variability is something that does not change—weight per unit size. The chaos that we would experience using paper would be unmanageable without invariance although most of us hardly know it exists. Scientists have gone to great lengths to establish invariants for dealing with the little things of life. In contrast, modern-day computer-oriented hearing scientists want us to believe that they can resolve the more important issues, such as brain functions that control our behaviours, through blind chance, effortless computations and relative judgements without invariance. To bring invariance into hearing research, let us first consider the procedure for quantitative evaluation of sensation, and then transition to neurophysiology of perception.

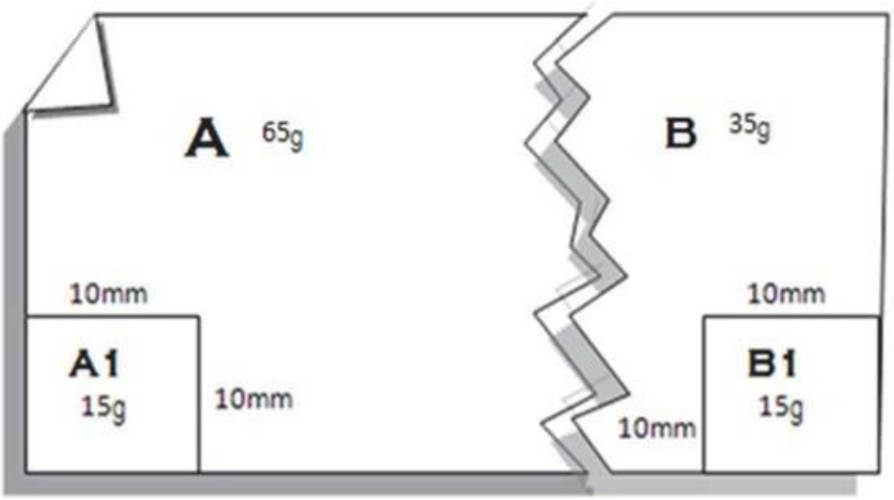


Figure 1: Invariance in variability. Portion A (65g) and Portion B (35g) of the same sheet of paper have precisely the same weight. (Image from Essien[11(p55)])

2.2. PHYSICAL CORRELATE

The quantitative estimation of any phenomenon at all requires a yardstick. To fulfil its purpose, the *relationship* between a yardstick and magnitudes of the phenomenon it measures *must not* change. Concerning quantitative estimation of sensation, there arose a problem: Psychologists about the 19th century reportedly argued that activities of the *mind lacked magnitude* and therefore were not subject to quantitative estimation.[15,16]. The fact is that we have never ever measured anything. Rather, we establish a yardstick and arrange units of our yardstick to reflect units of the targeted phenomenon. The success of this procedure is conditional upon invariance, that is, the relationship between units of the yardstick and units of the inaccessible phenomenon must not change That is the theory[15,16,17]. Let us consider how this works in practice to quantify sensation.

Temperature, for example, is intangible; it is inaccessible; it lacks mass, but we measure it thanks to the psychophysical principle of invariance. In this regard, fluctuations in the level of mercury in a thermometer correspond to fluctuations in temperature; by giving numbers to different levels of mercury in a tube, we end up with numbers that indicate units of temperature (hot/cold) without measuring temperature itself since it is inaccessible and intangible. It works. Although everything is changing, the functional relationship does not change because one cannot change environmental temperature and yet hold the level of mercury constant; nor can the level of mercury change while environmental temperature is constant. Thus, we measure temperature throughout the universe without disputes over linearity/non-linearity of the organism, logarithmic, and/or cyclical scales—controversies that characterize musical pitch and speech perception research in the absence of a physical correlate. That brings us to the stimulus. Is a physical correlate a stimulus?

2.3. THE STIMULUS

The physical correlate that is used to quantify a sensation is not necessarily the stimulus that the nervous system responds to when it experiences a sensation. For example, whereas fluctuations in the level of mercury in a thermometer permit the quantification of temperature, mercury or its fluctuations in a thermometer are not the cause of heat, nor are they accessible to the organism. Therefore, it would be unprofitable if a neurophysiologist smeared mercury on the skin of experimental subjects in the effort to discover how the brain measures temperature. In other words, the stimulus for a sensation must not only be accessible to the organism but also the cause of the sensation.

Again, we do not know such a parameter for pitch[6,8,9,12,18]. Indeed, to function as a science of auditory behaviour, hearing research is in need of a scientific foundation. To that end, we need a close look at the concept of sound.

2.4. ECOLOGICAL CONCEPT OF SOUND

Sound is the raw material in the study of music, speech, and auditory perception. Hearing research has attempted to survive on the physical definition of sound as vibrations in the media[19,20,21]. From the ecological standpoint, the sound source is much more than just a tool for creating pressure and vibrations in the media. A sound source can generate a sound quality only if it possesses the mechanical property that underlies that auditory sensation regardless of the pressure and vibrations it creates in the media. Therefore, we cannot set just any object in vibration to produce a musical or speech code. Rather, we *tune* the body, not to vibrate at a specific frequency, but to endow the source with the desired mechanical property for the code. The tuning process is the means by which a music performer, whether he knows it or not, endows an acoustic system with the code to transmit. This procedure is universal and applies in speech and music. Thus, this study considers *sound* as another form of its source, much like steam is another form of water or ice. Ice or water cannot propagate in the air, but if heated the resulting steam can. In like manner, a sound source cannot propagate in the medium, but if excited, the resulting sound—an intangible form of the source—propagates in the medium, conveying all the information about the source. In other words, when we hear sounds, we are in reality *seeing* and analysing physical bodies with our ears. In fact, the ear *sees* certain characteristics of objects through the sound much better than the eye does[2,22,23]. The definition of sound as the image of the sound source establishes the missing link between the sound source and the organism[24].This ecological conception of sound is the platform for this study.

2.5. RESEARCH HYPOTHESIS

The above brief presentation of psychophysical principle for dealing with sensory activities betrays the futility of 2,500 years of efforts to quantify pitch/tone without a physical correlate, and to explain pitch perception without the stimulus to pitch/tone. Therefore, there can be no valid hearing theory anywhere in the world without provisions that satisfy psychophysical requirements for scientific studies on sound and its attributes. The focus here is pitch. In this regard, the Pythagorean mechanical contribution, though the most plausible in hearing research, is considered here as an incomplete work in the absence of a mechanical invariant. Thus, the experiments in this paper take off where Pythagoras left off. The research hypothesis is based on the ecological conception of sound, guided by the concept of invariance as laid out above here: A sound source cannot transmit a code that it does not possess. The presence/ absence of a mechanical property determines the presence/absence of the auditory code in the sound. Wherefore, invariance in auditory perception begins at the sound production level. If a sound source transmitted only one information, the investigator's task would be relatively easier. However, every excited sound source transmits simultaneously several auditory attributes as a function of its physical, mechanical, and other properties. By implication, every auditory code is attributable to one (or more) mechanical property of the sound source. To master the principle of the auditory mechanism, we must identify the mechanical property of the sound source that underlies an auditory sensation, regardless of vibrational frequency which is the joint product of all the physical, mechanical, chemical, and other properties of the sound source.

Consider figure 2. It presents a drum (membrane), a guitar (strings), a xylophone (bars/ tongues), a trumpet (woodwind), a mouth (vocal tract). These instruments represent the five main categories of sound sources used in music and speech production. Interestingly, for these instruments to play harmoniously, all of them may be tuned to the *same* pitch, for example, the A440 tone. For the eye, all the instruments are different, but for invariance, all the instruments are the same because they are governed by the same law whereby *X*, a commonality of all instruments, underlies pitch *P* regardless of physical and mechanical variability of the sound sources, or of vibrational frequency which is determined jointly by all the properties of the entire sound source. The parameter *X* is the mechanical invariant that each of the instruments *must* possess to generate pitch. By implication, different quantities of *X* determine different pitch heights. Therefore, to understand pitch, measure pitch, create a pitch scale, establish pitch intervals, and unveil neurophysiological process(es) in pitch perception, we *must* discover the parameter *X*. This study provides experimental evidence for the identity of *X* from two acoustic systems—the Yoruba drum and musical strings.

2.6. EXPERIMENT 1: PERCEPTION OF YORUBA TONES

2.6.1. Yoruba tones

Yoruba is a three-level tone language spoken in Nigeria. Its three phonemic pitch levels (or tones)—Low. Mid, and High—provoke lexical contrasts in otherwise identical words. Outside this lexical role, pitch also serves intonational purposes, (a) to state a fact, (b) ask a question;

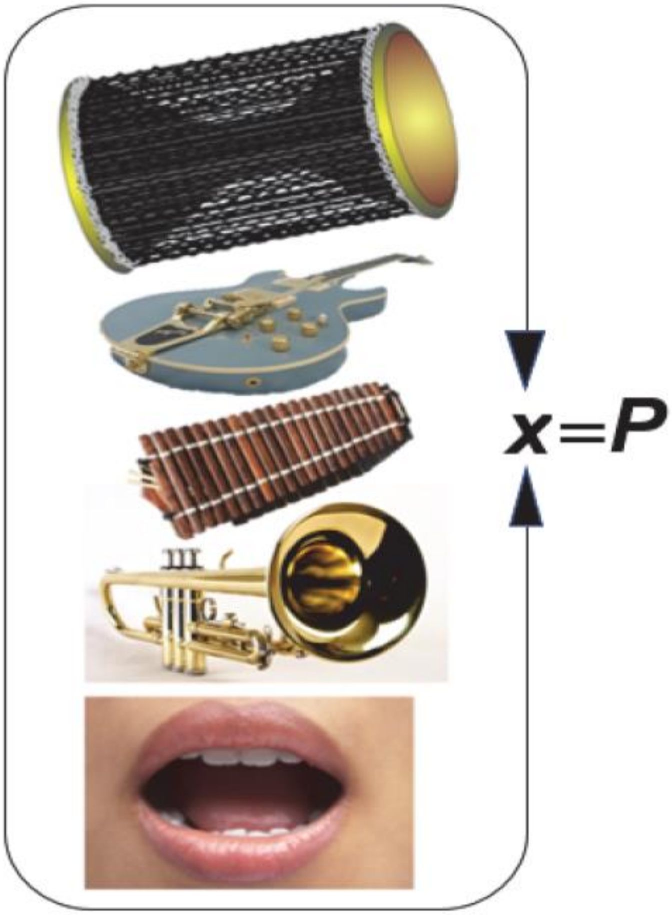


Figure 2: Mechanical invariance in pitch control. The instruments have different physical and mechanical characteristics. Yet, for invariance, they are the same thing in pitch production. All the different operations are different ways to adjust the same parameter *X* that controls pitch. (Image from Essien[25(p66)]

and (c) express emotions—joy, anger, sadness, surprise, etc. Besides, Yoruba songs and poetry exploit pitch modulation to create melody and rhymes, respectively. The multi-various roles of pitch in verbal communication in Yoruba (and other tone languages) raise long-standing yet unanswered crucial questions: How does the brain reduce all the audible pitches in nature to only three tones in Yoruba? How does the brain know when a pitch is lexical and when it is intonational? If pitch commutations provoke lexical contrasts, what becomes of lexical and semantic contents of expressions when pitch changes serve intonational and other expressive purposes? More complexity came with the introduction of instruments into the study of speech and music based on the vibrational frequency of the sound source. A Low tone at the beginning of an utterance may be higher in frequency than the High tone at the end; the High tone by a male adult may be lower in frequency than the Low tone by a female adult, whose High tone may be lower in frequency than the Low tone of a child. Perceptual constancy regardless of never-ending inter- and intra-speaker variability in tone production reduces the pitch/tone complexity to an enigma. And yet, the hearing scientist insists that frequency is pitch as it is tone. Thus, we know little or nothing about the way the brain handles the complexities speedily and effortlessly without disrupting verbal communication.

The auditory stimuli for the different perceptual settings mentioned above may be produced on the Yoruba drum (figure 2, top image), and are understood by Yoruba listeners; it is safe to say that all the secrets of pitch/tone production and perception are carefully concealed in the Yoruba drum. The experiments reported here below were designed to offer some insight into this complex human behaviour. To validate the choice of the Yoruba drum for this study, let us acquaint ourselves with some physical and mechanical characteristics of this instrument.

The Yoruba drum is a historic, compact, yet versatile acoustic system. Unlike traditional drums, it possesses a dynamic pitch regulating mechanism. Figure 3(a) illustrates the drum's hourglass-shaped wooden resonator. Each resonator head is covered with a skin membrane as shown in (b); the two membranes are connected with tendons as illustrated in (c), and the drum is complete, ready to play. The tendons may be squeezed toward the resonator to regulate the tension of the membrane; the more the tendons approach the resonator the higher the membrane tension and its pitch, and *vice versa*. The membrane is excited using the striker (d). Because the drum can produce also rising and falling glissandos, it can mimic any intonational pattern in verbal communication. This drum is most appropriate for the present study[25,26].

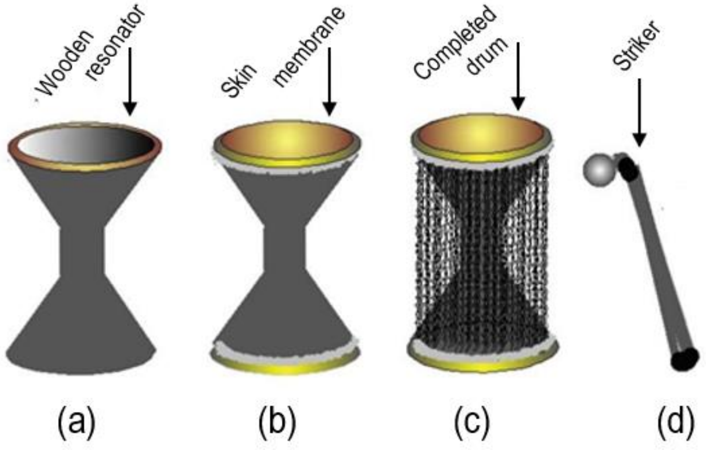
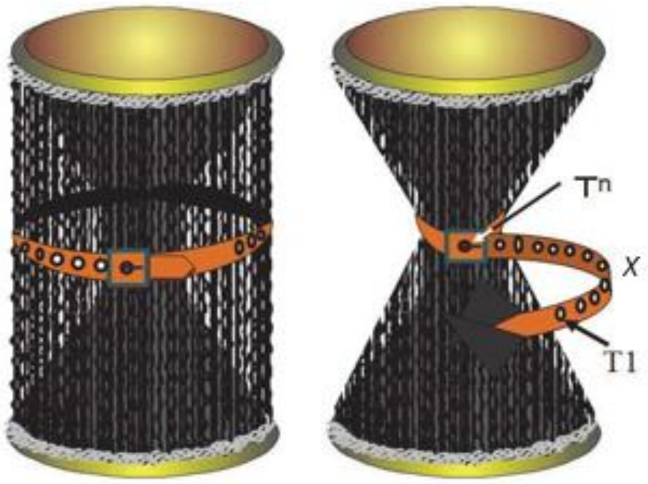


Figure 3: The Yoruba 'talking' drum. (a) The double-headed hourglass-shaped wooden resonator. (b) The drum heads are covered with skin membrane. (c) The membranes are connected by tendons. (d) The striker. (Artwork by the author, adapted from Essien[25 (p 245)]

2.6.2. Stimuli and test subjects

The stimuli for this experiment were acquired as follows: To obtain a physical representation of membrane tension, a belt was slipped around the drum's tendons. The belt has no effect on the membrane in the position illustrated in figure 4(a). The belt was then tightened around the tendons, squeezing them inward until they came up against the waist of the resonator, thus stretching the membrane to the maximum possible extent. A portion (x) of the belt was displaced as illustrated in figure 4(b). The displaced portion was notched every 15 mm. Each notch represents a different degree of tension and a different pitch. The signal generated by the membrane in the position illustrated in figure 4(a) was labelled T1 (for Tension 1). Then, the belt was tightened and arrested at the following notch, and the signal produced was labelled T2, etc. Sixteen signals (T1 to T16) were produced. The procedure ensured that the tension and the pitch at T1 were lower than at T2; and in turn, the tension and the pitch at T2 were lower than at T3, etc. Thus the tension and the pitch of each recorded signal correlated with each other, varying only as a function of membrane tension since diameter and thickness were held constant.

Twelve subjects (hereafter Ss) took this test. Only one subject was a Yoruba linguist/tonologist; the other 11 Ss were linguistically naïve native Yoruba speakers; they received no prior training whatsoever, but because they were well acquainted with drum speech, they were informed that a Yoruba drummer played the message /ó lɔ sùn/ (he/she goes to sleep) on his drum; and that the recorded signals were accidentally mixed up. Their task was to identify each drum signal either as the (High-toned) word /ó/, or the (Mid-toned) word /lɔ/, or the (Low-toned) word /sùn/. The questionnaire offered three choices (ó or lɔ or sùn) for each signal. This test method was very effective because linguistically naïve Yoruba natives talk but know little or nothing of the role tone plays in the language. The signals were presented over loudspeakers in six different randomized stimuli sets, and three ascending and three descending orders of tension/pitch.



(a) (b)

Figure 4: Graduation of drum membrane tension. This method was devised to have the tension of experimental drum membrane under control. Illustration from Essien[25(p260)].

2.6.3. Results of Test 1

Figure 5 presents the summary of the test results. As a classificatory rule, each signal had to attract 60% correct tone labels to enter into a tone category. Figure 5 (top image) compares the performances of Ss in the randomized (R%) and non randomized (R/NR%) stimuli sets. Figure 5 (bottom image) compares the performances in the ascending/descending (A/D) orders of tension/pitch. These results show that despite errors each signal preserved its tone quality in every pitch environment. The observable improvements in performances resulted from the fact that some naïve Ss quickly acquired some experience with tone and this was manifest in their performance in the course of the experiments. Five tone categories were manifest: (1) Low (comprising signals T1 through T3); (2) Low/Mid (T4 and T5), (3) Mid, (T6 through T8); (4) Mid/High (T9 through T11); and (5) High (T11 through T16). Outside errors of judgment on the part of some naïve Ss, the pitches did not change their tone categories in any pitch environment. Very detailed analysis, applications and deductions in terms of tone perception theories, tone distances, the co-existence of tone and intonation in a tone language, up-step, down-step, downdrift (or terracing), etc, are discussed in detail elsewhere[10,16]. In this study, these results provide a stepping stone to our goal—the identity of the parameter *X* in figure 2.

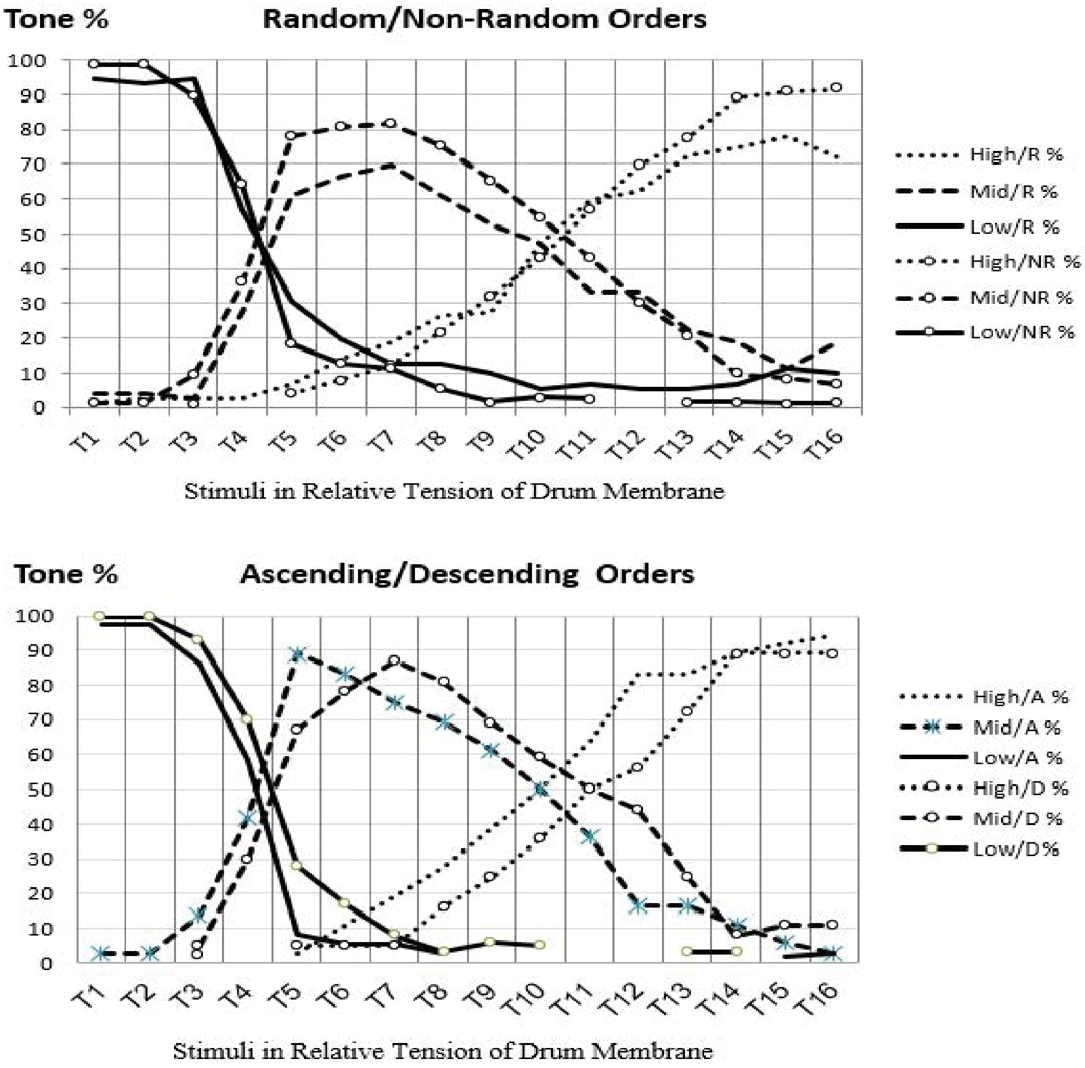


Figure 5: Tone qualities of drum pitches in different pitch environments. (Image from Essien[25(p274)]).

2.7 TEST 2: THE RESISTANCE OF TONE ZONES

The pitches in figure 5 preserved their tone categories in any pitch environment in contrast to the claim of relative pitch hypothesis of tone. To test the resistance of tone, and the resistance of the observed tone zones, another hypothesis was formulated: If tones derive their identities from relative judgments arising from comparisons of contiguous pitch heights, all the pitches in the stimuli set will redistribute into Low, Mid, and High tone categories if a tone zone were removed from the complete stimuli set.

2.7.1. Stimuli, Test Subjects, and Test.

In accordance with the hypothesis for this test, the Low tone zone comprising stimuli T1 through T3, which was the best perceived tone category, along with signals T4 and T5 in the Low/Mid transition zone, were removed from the stimuli paradigm. Stimuli T6 through T16 were presented to the same 12 Ss in 6 random orders following the same procedure as the previous test. The questionnaire was adjusted accordingly.

2.7.2. Results of Test 2

Figure 6 presents the summary of performances in Test 2. It compares the tone qualities of signals in the complete stimuli set (LMH) with the tone qualities of the same signals in the MH stimuli set following the removal of the Low tone zone. We observe increased performances in the identification of the Mid and the High tones. As noted earlier, the Ss acquired improved consciousness of tone. Of interest is the fact that no pitch in this test qualified to enter into the Low tone category (60% tone score). The Low tone quality of signal T6 at the Low/Mid tone boundary improved from 20% to 40%, yet it was better identified as a Mid rather than Low tone. Thus, despite improved performances, the highlight is the absence of the Low tone zone; each signal preserved its tone category in every pitch environment. Against expectations, the pitches failed to redistribute into three tone categories although the Ss had three choices.

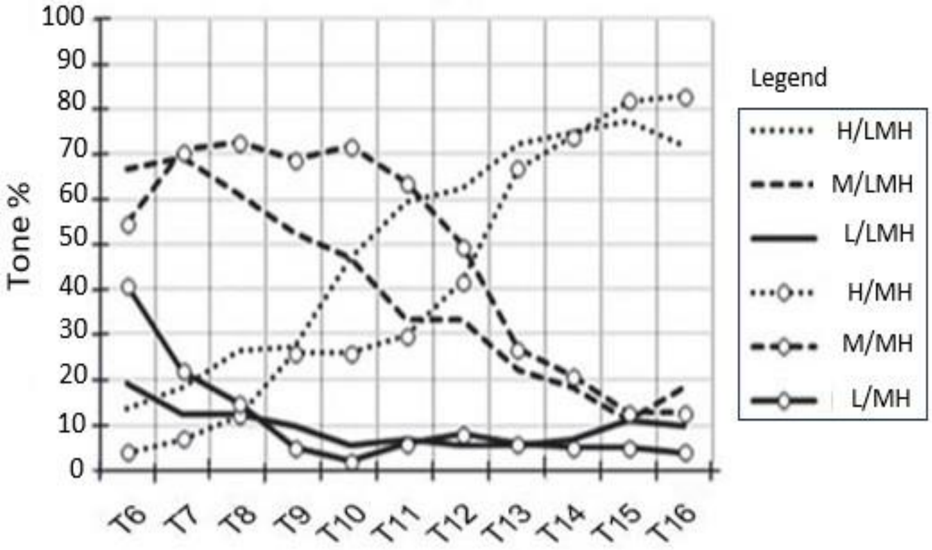


Figure 6: The resistance of tone qualities of signals in different pitch environments. Image from Essien[25(p283)]

2.8. OBSERVATIONS

Hearing scientists, even linguists, speak of tones in tone languages as though tone language users produce and use only two or three pitches in the pitch repertory. In contrast, the above test results show that a phonemic tone comprises several pitches in a tone zone. The benefits of this multi-pitch structure of a phonemic tone, the flexibility, and the freedom it offers in verbal communication are discussed in detail elsewhere[26,27]. Above all, the performances testify to the psychological reality of pitch and tone, perceived on absolute judgment basis without the need for pitch *Y* to perceive pitch *Z*. And if we needed *Y* to perceive *Z*, how did we perceive *Y*? By this evidence, pitch and tone are subjectable to quantification. These observations encourage the search for the hypothetical invariant parameter *X* in figure 2.

2.9. ON MECHANICAL INVARIANCE IN TONE PERCEPTION

2.9.1. Toward mechanical invariance in pitch production

If all sound sources produced pitch and tone according to the data we have just examined, we could confidently conclude that the parameter *X* is the force applied to the membrane. For invariance, the observation would be gratuitous. Why? Outside the balanced force applied, another variable that impacts pitch is the membrane diameter, and in the case of a string, its length and density. Therefore, the observable relationship between the balanced force and pitch in the above experiment might break down and send pitch, tone, and tension their separate ways if another parameter outside tension was modified. The question arose whether experimental data from other drums would match those on hand. Thus, other drums were brought in to satisfy the requirements of invariance.

Figure 7 presents three drums in three different sizes as indicated. The relevant parameter in this experiment is the diameter which measured, from the largest drum to the smallest, 230mm, 145mm, and 105mm. The largest Drum A was used in the preceding experiment.

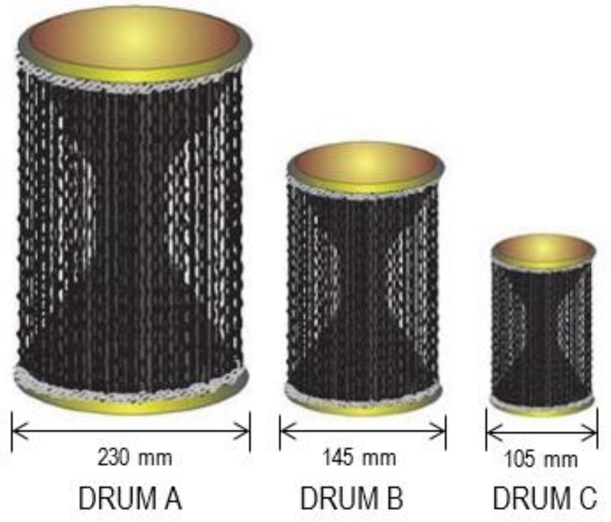


Figure 7:Three experimental drums in different diameters

The stimuli from the other two drums B and C were acquired following the same procedure described earlier (see figure 4). Sixteen signals were produced on Drum A for the preceding test; ten on Drum B, and eight on Drum C. The same 12 test subjects (Ss) took the present tone labelling tests under the same conditions described earlier. Detailed test results, inter drum variability, inter and intra-subject variability are described and discussed elsewhere[25,26]. The relevant highlight for the present study is presented here below.

Figure 8 compares tone qualities of pitches produced by the three drums. The highlight is the Low tone quality of the pitches produced by the three drums. The largest Drum A produced the best quality Low tone pitches (94%). However, the lowest pitch from Drum B was 76.5% Low. In contrast, the lowest pitch from Drum C was 10% Low. We observe that as drum size decreases, it loses the capacity to produce the sensation of lowness in pitch.

At first glance, following research trend in psychoacoustics, one could jump to the conclusion that the loss of the Low tone results from increased frequency as drum size decreases. We shall examine that possibility later below here to verify contrary opinions. In the meantime, it is necessary to recall that the present study aims to provide the missing mechanical invariant so that we have a firm hold on what a sound source transmits as the code for pitch. Besides, vibrational frequency should not come into the discussion because it is *not* a parameter of the sound source. Therefore, we must assume that this study is conducted in the 6th century B.C., and that we have no knowledge whatsoever of frequency of vibration, Ohm's acoustic law, Helmholtz's resonance/place theory, nor Fourier analysis. All that we have are three drums, and it is known that the smallest drum does not produce Low tone signals[28]. The mechanical orientation of this study raised the question: How does size deprive Drum C of the capacity to produce Low tone pitches? Another hypothesis for another experiment was formulated to discover the mechanical origin of the missing Low tone and help build on what we had already acquired.

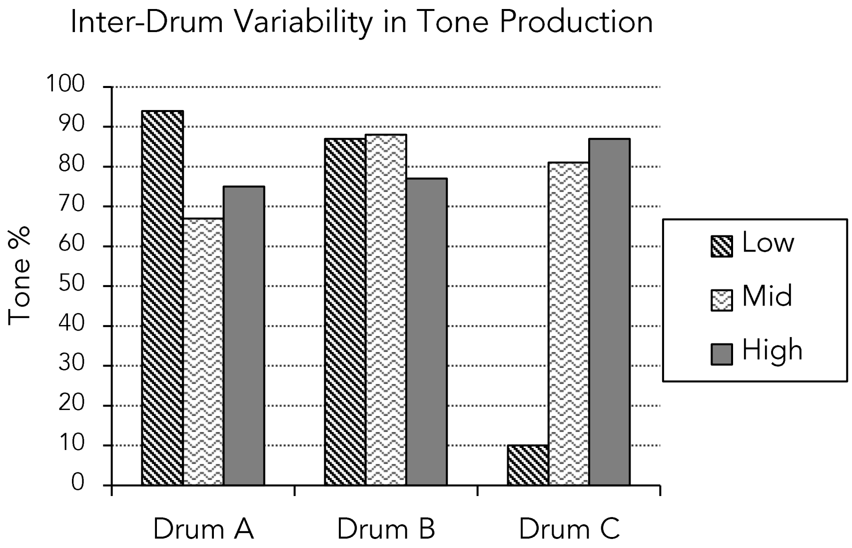


Figure 8: Inter-drum variability in the production of Yoruba Low, Mid, and High tones. The Low tone quality of drum signals decreases with drum size, culminating in the total loss of the Low tone quality in the signals generated by Drum C. Figure from Essien[25(p293)]

2.9.2. Test 3: Pitch Equivalents Across Drums

The pitches produced on the membranes of drums A, B and C maintain a functional relationship with membrane tension. However, the pitches of one drum are not necessarily the same as those of another drum. Since reduction in drum size shifted the lower threshold of pitches produced by the relatively smaller drum up the pitch scale, this experiment aimed to determine the mechanical factor that underlies the pitch shift and the degree of shift. Accordingly, to explain the loss of the Low tone quality in terms of a property of the sound source, a pitch perception test was organized to establish pitch equivalents across the three experimental drums A, B, and C.

2.9.2.1. Acquisition of Stimuli

The stimuli for this test comprised all the signals produced on the three drums. A playlist shown in figure 9 was established and fed to a computer using the software *Sound Designer II*. According to the playlist, each time the computer played the signal BT1, it was followed by a pitch from Drum A. The arrangement produced an AX paradigm with signals paired up as follows: BT1/AT1, BT1/AT2, BT1/AT3 ... BT1/AT16. This structure of the playlist makes it possible to match the pitch of BT1 with that of a signal from Drum A. Stimuli pairs were separated by 500 milliseconds of silence. The procedure was repeated using pitches produced on the three drums. The drums were paired up as illustrated in figure 9. Also, another stimuli set was formed to establish the acoustic capacities of drums: In these stimuli pairs, the highest pitches were paired up as follows: BT10/AT16; CT8/AT16; and CT8/BT10.

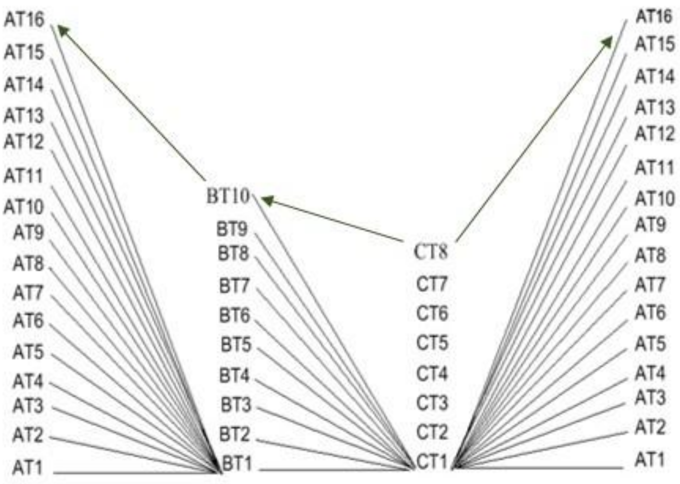


Figure 9: Structure of pitch-matching test to establish pitch equivalents across drums A, B, and C. Image from Essien [25(p316)]

2.9.2.2. Test 3: Subjects and Test

The nature of this test restricted participation to only persons who were well acquainted with pitch and the requirements of the test. Eight doctoral research students at the Institute of Phonetics, Sorbonne University (Paris III) participated in the test. The test subjects (hereafter *S*s) had each a questionnaire in which paired stimuli had been listed following the order of the stimuli in the playlist. Their task in the first part of the test consisted in detecting the signal pair with equal pitch. In the second part of the test, the *S*s had to indicate whether the first signal was lower, equal, or higher in pitch than the second. The *S*s could pause the recording at any point in the test, and repeat it at will as many times as they needed to establish pitch matches or judge relative heights of stimuli pairs. The stimuli were presented over hi-fi quality headphones. Only one subject was tested at a time.

2.9.2.3. Results of Test 3

The test results are summarized in figure 10 showing pitch matches for the three drum pairs B/A, C/A, and C/B. The results for the B/A stimuli pairs show that 75% of equal pitch labels were assigned to the stimuli pair BT1/AT4. The remaining 25% was shared equally between AT5 and AT6. In the C/A stimuli pairs, 75% of equal pitch labels were assigned to the pair CT1/AT7; the remaining 25% went to AT6. In the C/B stimuli set, all *S*s unanimously assigned equal pitch labels to the pair CT1/BT6. In Part 2 of the test, the results show unanimous responses from the *Ss* that BT10 was higher in pitch than AT16, CT8 was higher in pitch than AT16; and the same CT8 was higher in pitch than BT10. Let us now examine these results closely in the effort to resolve the mystery of the missing Low tone and the pitch slide up the pitch scale.

2.10. SOUND PRODUCTION AND PITCH PERCEPTION

2.10.1. Discrepancies in hearing research

The results of the pitch-matching test throw light on otherwise unexplainable issues, particularly the case of the missing Low tone, and the mechanical origin of pitch. Regarding the missing Low tone, we find that BT1 and AT4 have the same subjective pitch; it is totally natural that the Low tone quality of signal BT1 manifested the same poor Low tone quality as signal AT4 even though BT1 is the lowest pitch of Drum B. Similarly, because CT1 and AT7 generate the same pitch, and the two signals were labelled a Mid tone in every pitch environment, Drum C had lost the capacity to produce a Low tone even though CT1 was its lowest pitch. The evidence suggests that the ear measures a property of the sound source to allocate a pitch to a tone category. To identify the property, let us reason from the production data in figure 10.

Consider the first drum pair A and B in figure 10. The illustration shows that to produce signal AT4, a certain amount of balanced force was applied to the membrane to raise the drum membrane tension from AT1 to AT4. Yet, Drum B produced the *same* pitch without the need for a balanced force. The second drum pair is A and C. Drum A required a certain amount of balanced force to raise the membrane tension from AT1 to produce signal AT7. However, Drum C produced the *same* subjective pitch without the need for balanced force. The same is true of the drum pair B and C.

The above facts raise a question: Why does one drum require balanced force to produce a pitch and another drum does it without the need for force? We must resolve this psycho-mechanical relationship to find the parameter *X* in figure 2. In this regard, Newton's law states that a body cannot increase its velocity unless the force exerted on it increases[29]. A question arises: When the diameter of a drum is reduced, and the membrane increases its vibrational frequency to produce a higher pitch without a balanced force, where does the increased force that drives the membrane come from? Apparently, the answer for this question has the potential to demolish and cart away the entire acoustic foundation of hearing research. Let us see.

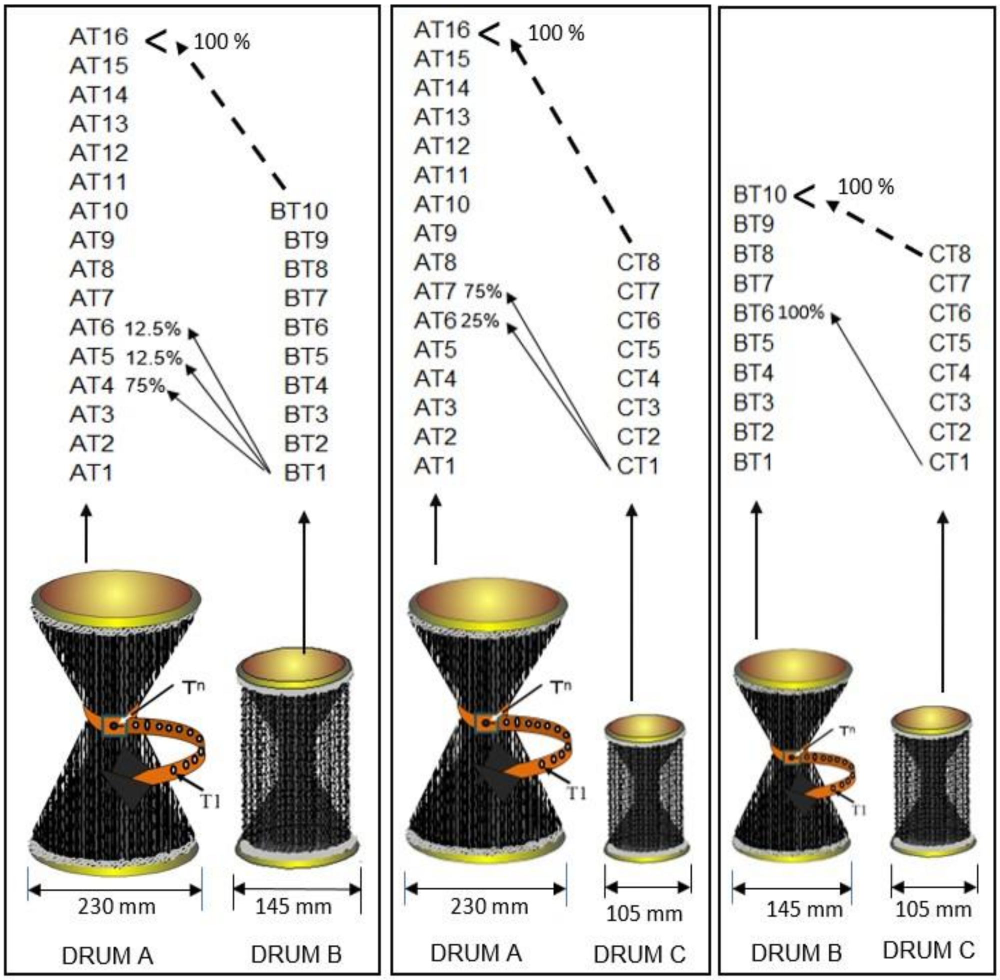


Figure 10: Pitch equivalents across drums, showing the interaction between the force exerted on the membrane and the diameter of the membrane in pitch production. When the diameter of the membrane is reduced, where does the force that accelerates the membrane come from?

We have been taught and made to believe that the force in a string is the balanced force exerted on a string, and that the force is constant as long as the balanced force is held constant. The provision cannot explain the above experiences with drums. Is the situation any different with strings? No. When a string is shortened, it accelerates and produces a higher pitch although the so-called tension is held constant. Where, then, does the extra force that drives the string at a higher frequency come from? And how can *increasing* rates of vibrating sub-lengths of a string be proportional to the square root of the tension that is held *constant*? Is the error with Newton's law? Or is it with pre-historic physics of sound in hearing research? Many, many more unanswered questions [30] expose hearing research as standing on a scientifically unfounded foundation. [31] And answers to the above questions fall outside the provisions for the drum experiments in the absence of a system for measuring the force exerted on the membranes.

To address the issues, a further experiment which is reported in detail elsewhere was conducted on strings[25,31]. The summary of the experiment is presented here below to conclude this investigation on the primacy of sound sources in pitch/tone perception.

2.10.2.Test 4: String, Sub-lengths, and Pitch

2.10.2.1. Test 4: Equipment

The data in figure 10 call for terminological distinction between the force applied externally to a membrane (called tension) and the force in the membrane. To address this situation, a special device shown and described in figure 11 was designed by this author for construction by a guitar manufacturer in Reading, England, United Kingdom. Among its many uses, the one experiment that directly addresses the present problem is reported here.

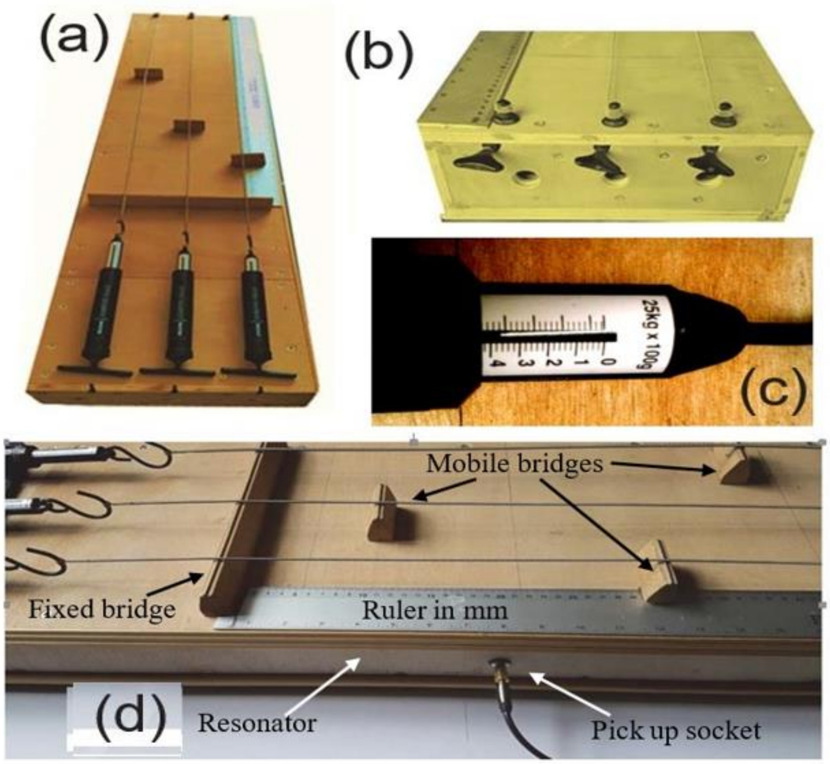


Figure 11: The multi-cord tuner. (a) The complete device with three strings in different densities under tension. (b) The balanced force exerted on each string is adjustable via tuning heads. (c) The force is recorded in kg by a spring balance for each string. (d) A ruler graduated in mm permits reading off the effective length of string at the position of each mobile bridge. An integrated pick-up socket allows for capturing the tones produced using a sound recorder. (Image adapted from Essien[25(p365)]

2.10.2.2. Stimuli, Test subjects and Test

The stimuli for this test were five musical tones of the same pitch, one was produced on a full string, and four on sub-lengths of the same string. The stimuli were produced as follows: An A85 musical string was mounted on the tuner in figure 11. In compliance with the requirements of invariance, the full string and its sub-lengths must be tuned to sound the same pitch. Two Ss, a male and a female, took part in this task. The male participant had played guitar for over 5 years with a group on part-time basis. The female, who possessed the rare ability of absolute pitch, was a professional classical violinist with an M.A. degree from the Royal Academy of Music, London, UK, lecturer in classical violin at Pimlico Academy, London, UK; she had played violin for 23 years from infancy. Their task was to tune the A85 string and its sub-lengths to the same musical pitch A220. The string tension was preset at 22 kg. One S took the test at a time. The male S could not complete the task when the string got too short, but what he could do agreed totally with that of the female S who had no problem whatsoever.

2.10.2.3. Results of Test 4

Figure 12 is a summary of the results. At the 22 kg. preset force, the effective string length was 860 mm. Thereafter, string length was reduced in steps to 645 mm, 430 mm, 215 mm, 107 mm, and 53 mm. The figure shows that the 645-mm string required 13.2 kg force to sound the *same* pitchA220. Thus, by shortening the string, the 645-mm-long string acquired 8.8 kg. force which had to be displaced to lower its pitch to A220. The figure shows that as string length decreased, the force in the string increased, and more units of force had to be displaced for subsequent sub-lengths to sound the same pitch. The 53-mm portion required no force at all from outside.

2.10.2.4. On the Results of Test 4

The above experiment brings the mechanical investigations in this study to its conclusion. We deduce from the results that the force in a string (hereafter *F*in) is *not* constant when string length is reduced even though the force exerted externally (hereafter *F*ex) is held constant. The experiment converts decreasing string length into *increasing force* to explain *increasing pitches* generated by sub-lengths of the string. Unlike string length, force maintains a direct and functional relationship with pitch. This finding redresses a fundamental error in hearing research. It brings the string ratio theory to completion by introducing invariance into its variables to deal scientifically with pitch, tone, and auditory perception as a whole.

Now that we have a mechanical parameter that maintains a direct and functional relationship with pitch/tone, let us close this study by briefly examining the way the production parameter *force* impacts physical representations of sound and its perception

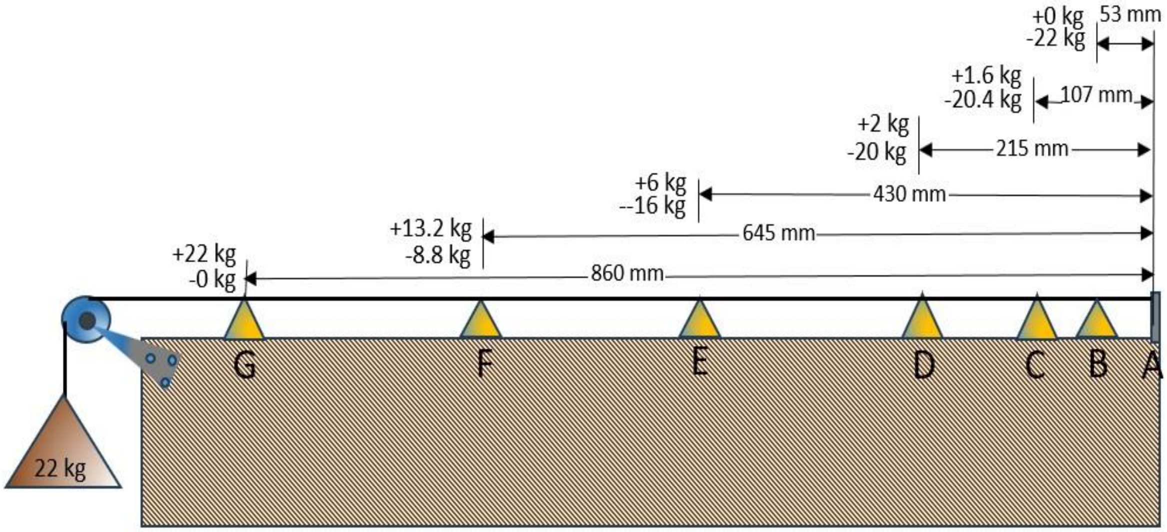


Figure 12. Two sources of force in a string. (1) The balanced force supplied externally (*F*ex). (2) The force that is the inherent property of the string (*F*in) which varies as a function of string length. Contrary to laws of physics and mathematics of sound and music, the balanced force (*F*ex) does not determine the force in the string.

2.11. FORCE, SPECTRAL STRUCTURE, AND PITCH

2.11.1. Drum signals

Figure 13 presents the waveforms and the spectral structures of two drum signals AT1 and CT8, representing the lowest and the highest pitches, respectively, in the drum experiments. The signal AT1 was produced without exerting force on the membrane except what was necessary for the membrane to vibrate.. The waveform of the slack membrane (left image) is irregular and short; rapid attenuation of amplitude follows the burst at onset of vibrations. The non-harmonic spectrum shows a high peak at onset; the rest of the signal is low intensity noise with rapid energy attenuation in the high frequency region. In contrast, signal CT8 (right image) is longer; it has a rich and regular waveform with high amplitude at onset of vibrations followed by gentle and gradual attenuation of energy to the end of the signal. The spectrum comprises only one Fourier component. In fact, all the signals produced on Drum C sounded like pure tones—a quality that is evident from the spectral structure. Yet, the two signals were produced on membranes which are said to produce indistinct tones. To explain this variability within the framework of ecological acoustics and invariance, let us examine the data from strings which produced one universally established auditory code.

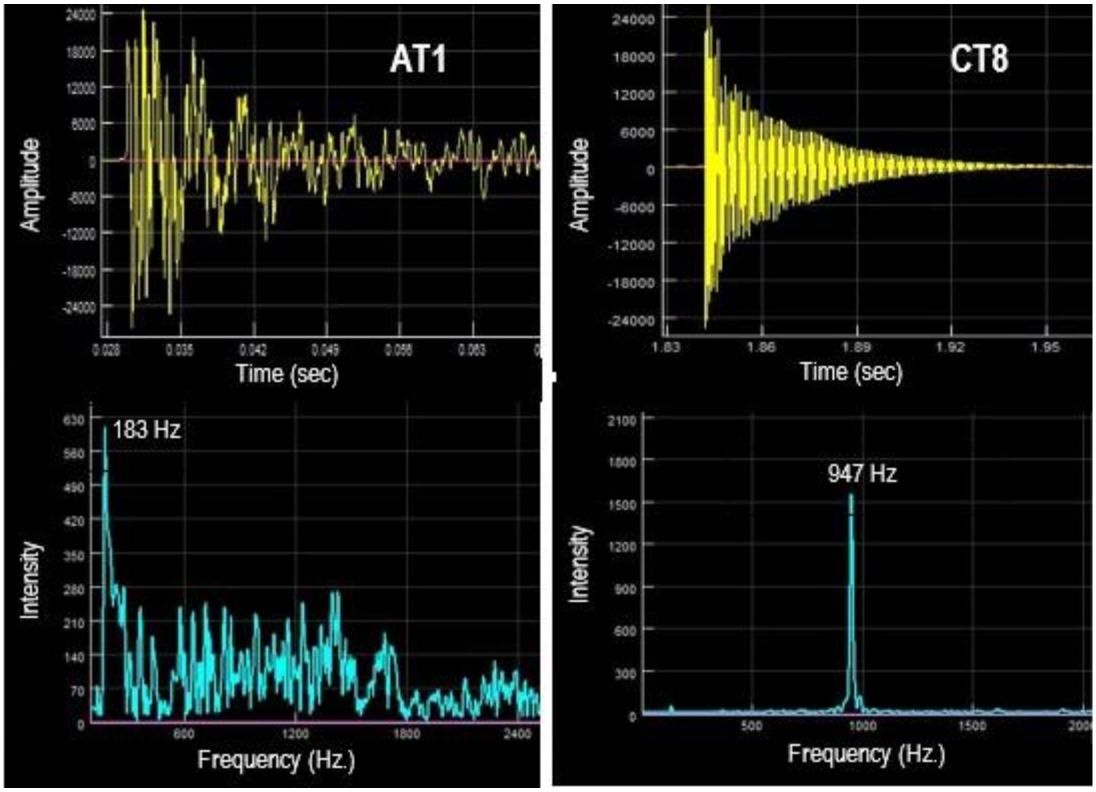


Figure 13. The impact of force in the sound source on acoustic representation of sound and the sensation pitch. The two signals above differ in pitch and in spectral structure as a result of their different mechanical configurations.

2.11.2. String signals

Figure 14 presents the waveform and frequency spectrum of the A220 musical tone produced on the 860-mm-long string. The signal, from onset of vibrations to the end, is 10.2 secs long. The waveform is squeezed to fit into the picture. The precision of its structure is very close to mathematical perfection as the inset picture shows. It preserves the peak-to-peak interval almost to the last cycle. This is reflected in the precision of the frequency spectrum throughout its 16 Fourier components numbered 1st through 16th.

However, the above signal defies all acoustic theories of musical pitch perception: The most prominent spectral component is the 165 Hz. partial; it is not the fundamental (hereafter f_o) but the third harmonic partial, showing that the f_o at 55 Hz. and the second partial at 110 Hz. are missing. Since the early 1920s, this phenomenon—the missing fundamental—has dominated hearing research to this day such that all "our so-called theories of hearing are actually only theories of pitch perception." [32(p589)] Yet, the problem is unresolved. Psychoacoustic theorists have attempted to explain the pitch of signals such as this by attributing pitch to one or another of four candidates: (1) The fundamental (f_o); (2) the position of maximum stimulation on the basilar membrane; (3) the 220 Hz. component, and (4) the 'residue'.

Concerning (1), the argument is refuted by the absence of the f_o from the signal. Critics insist that the auditory system cannot function with what is not accessible to it (see section 2.3

above)[33,34,35,36]. Then, proponents of the theory claim that the missing f_o is generated according to the beats hypothesis which involves the difference between the values of two contiguous harmonics. Again, the claim encounters a problem because the auditory system in the present case would retrieve 55 Hz. which cannot generate the A220 tone. Today, the equal pitch of signals at different octaves are attributed to pitch chroma. But since different octaves have different frequency values, pitch chroma fails to explain how different frequencies give rise to the same pitch. And the arguments go on. Concerning (2), the 165 Hz. component cannot be rightly expected to elicit the perception of the A220 tone. With regard to (3), the rightful candidate is the 220 Hz. component which is present in the signal. However, it is not the f_o, and because of its reduced energy, its stimulation on the basilar membrane cannot gain prominence over the 165 Hz. component. (4) When the f_o is absent or removed from the signal by masking, what is left of the signal as candidate number 4 is called 'the residue'. We shall consider this phenomenon in the discussion later.

Another problem is concealed in the very misleading nature of the above 2D presentations. Spectral slices and 3D visualizations (which are published elsewhere [25,30] but cannot be presented here for lack of space) show that the 220 Hz. component is present at low intensity levels only in the last three of six spectral slices of the signal. Any conclusion on brain functions in pitch perception drawn from the above presentation is meaningless because the data do not support any[5,7,18,25]. The positive point, though, is that the 220 Hz. component is in the signal.

What can the other signals offer? The same A220 musical tone was produced on the sub-lengths of the string. The signals of importance to us are those produced using sub-lengths that would, in theory, produce pitches at Octave intervals; i.e. 430 mm, 215 mm. 107 mm, and 53 mm. Their fine-surface structures are presented here below.

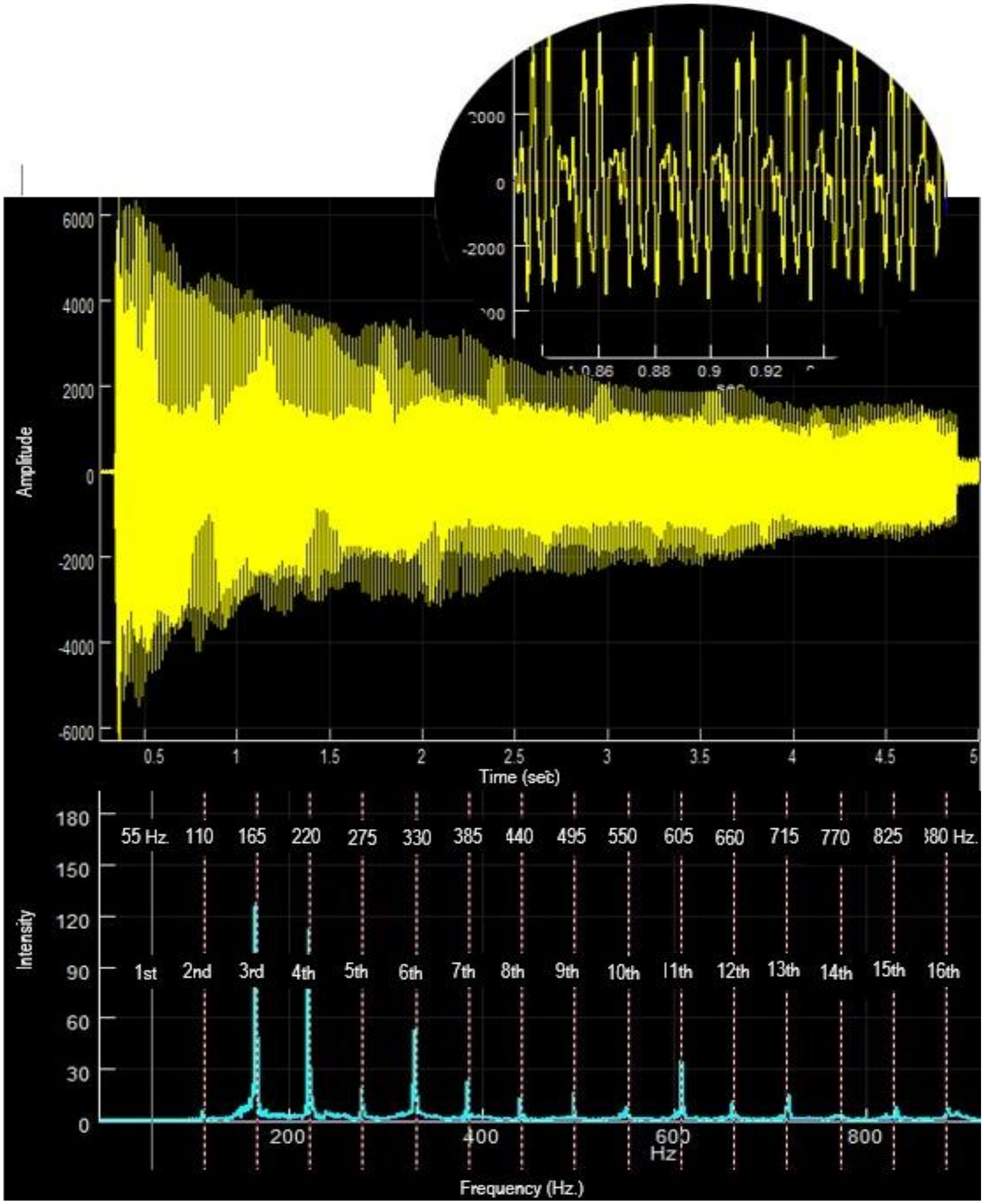


Figure 14: Acoustic data from the A220 musical tone produced on the 860-mm-long string, under the balanced force of 22kg

Figure 15 below presents the acoustic data for the same tone produced on sub-lengths of the string. In comparison to the robust, vibrant, and lush waveform of the 860-mm-long string, gradual impoverishment of the physical characteristics of signals occurred as string length decreased. The signal from the 430-mm-long string is shorter; its .resonant frequency stands at 164 Hz. with the f_o at 54.8 Hz. and three other partials with little or no energy. The first two harmonic partials are missing. Further reduction in string length, or alternatively, as the string's inherent force (*F*in) rose, the waveform for the 215-mm string turned skeletal and the spectrum anharmonic. The oscillograms and the spectral structures are similar to that of the lowest pitch on the largest Drum A. Despite the remarkable variability of the acoustic characteristics of the string signals, each and every one of them generated the same auditory musical code. Any efforts to explain perceptual constancy through the above irreconcilable acoustic variability can only lead to many more square pegs in round holes. That is precisely a fitting description of hearing research today, in theory and in practice, in the absence of invariance at the mechanical level for a guiding principle. Primacy must be accorded to the sound source for hearing research to attain the status of a science of auditory behaviour.

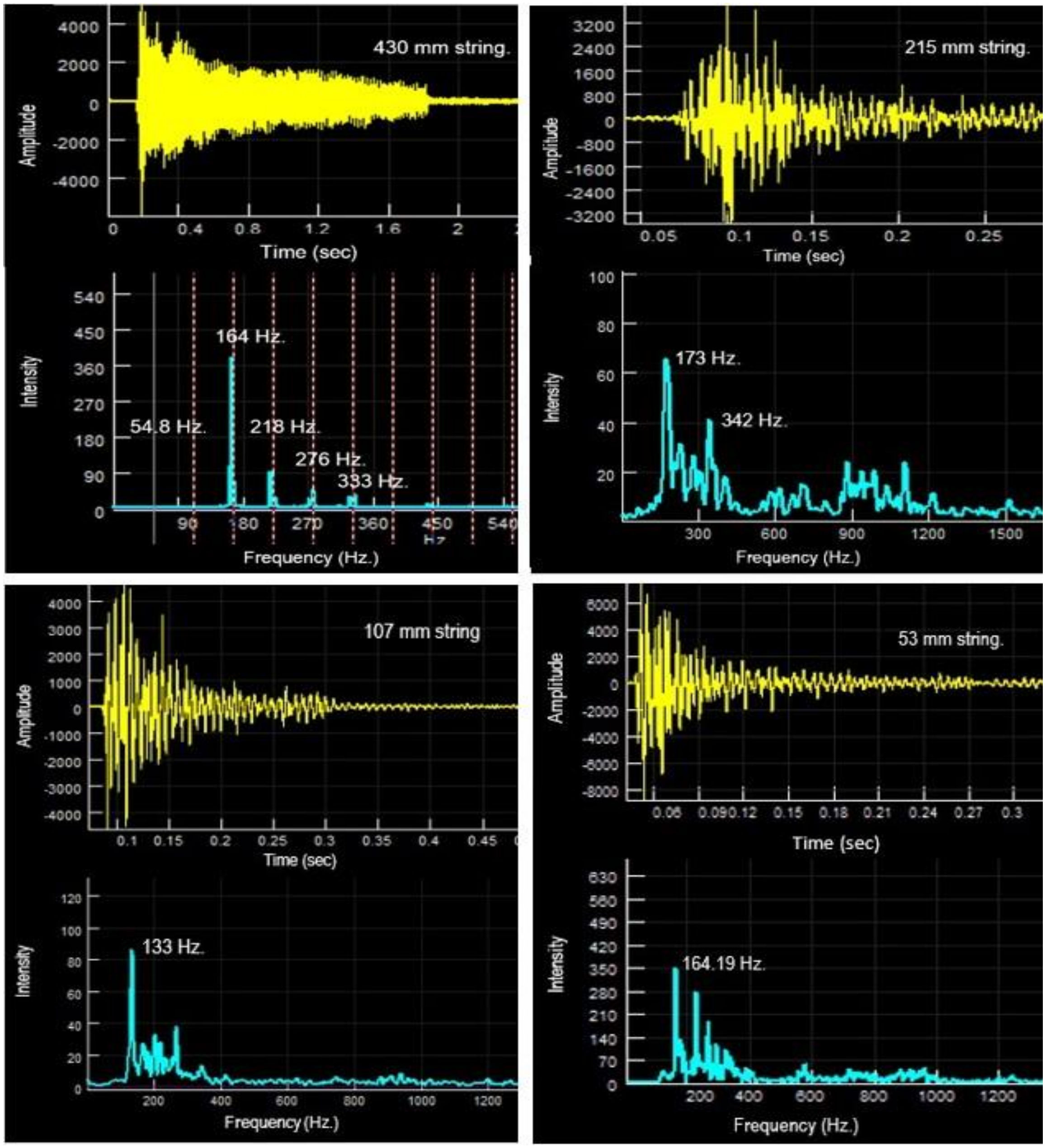


Figure 15: Impact of force on spectral structure and pitch. Acoustic data from the A220 musical tone produced on sub-lengths of the same string

3. Discussion.

Apparently, the way the brain handles pitch is the gateway to the secrets of auditory perception. Yet, the foremost challenge in hearing research is accurate understanding of pitch. The failure to explain pitch implies that hearing research has not taken a step beyond its starting point since 2,500 years ago. Nevertheless, the science had a perfect mechanical start. Where did it go astray? What readjustments are needed? This discussion will focus the contribution of this paper, and the implications for the future of pitch and auditory perception research.

For invariance, sound production, acoustics, and neurophysiology of perception are one and the same thing. After all, the message that leaves the source (production) and goes through the medium (transmission) must be received at the destination (perception) otherwise there is no communication. At the three levels, the focus is the message. Complexities arise at all three levels because the source transmits the intended message and many other pieces of information simultaneously through the medium. The researcher must identify in the flux and isolate the message that provokes a specific response in the receptor organism. It is very challenging. Nevertheless, the experiments in this paper prove that however hard it might seem, it can be done, but not without invariance.

Pythagorean string ratios, although a perfect start, comprised only variables. String length is not invariant with pitch. The conversion of string ratios into frequency ratios without invariance with pitch was not a remedy for the problem because frequency is not a parameter of the sound source. To erect a behavioural science on a variable that has no relationship with pitch was a fundamental error in hearing research, with profound impact on all aspects of hearing sciences. The practice attributes the source of the message to the medium for which reason the sound source is insignificant in auditory perception; the emphasis is on the pressure and wave propagating in the medium, and the sound source is only a means to an end. The practice ignores the fact that not all (if any) information-bearing elements in auditory perception have acoustic representations. Therefore. the scientist amasses a paraphernalia of sophisticated technology to study patterns of vibrating air particles and establishes computations with the hope to explain how the brain works without the least knowledge of what the brain responds to when stimulated with sound. The classic samples of sound wave examined in this study failed to support any acoustic theory of pitch. Yet, the hearing scientist believes (or wants everyone to believe) that spectrographic material and Fourier frequency spectrum are visualizations of how the ear hears sounds[37,38]. Let us consider some disparaging opinions of the practice to highlight the need for a better replacement.

Many linguists unequivocally rejected the new technology without compromise right from the outset[39]. For Mol, [40(p8)] acoustic approach to auditory perception is a “narrow-minded way of thinking which leads us into a blind alley when faced with the problem of the real nature of the auditory mechanism." It would be “quite healthy for engineers and kindred spirits,” Mol[41(p13)] pursued, “to come down to earth every now and then in order to realize how and if their mathematical concepts fit into the real mechanism of speech and hearing." Fant[42(p174)] described acoustic features in perception "as imaginary and empty as the 'Emperor's New Clothes.'" The case of the missing fundamental and the pitch of the residue (see above) was reserved for this discussion. In this regard, it has not happened to this author to produce a signal with the f_o in the spectrum. The position of the resonant frequency component in the spectrum is equally erratic. Boomsliter & Creel[43] replicated Helmholtz's siren experiment, but they failed to find the f_o in the signal. Radocy[30(p75)] observed that the problem is not the lack of a one-to-one relationship between frequency and pitch, but the "inconsistent inconsistency" in the relationship. Regarding the pitch of the residue, most investigators exploit synthetic signals. They attempt to explain pitch in terms of the so-called critical frequency (or dominance spectral region)[44-50]. Whereas the experimenter with naturally-produced signals has (very little or) no control over the spectral composition of the stimuli, the experimenter with synthetic sounds can construct stimuli to satisfy his/her spectral fancy. This leads to compatibility problems between natural and synthetic sounds. The validity of results obtained by the use of such signals which may not be produced by natural acoustic systems has been highly questionable and controversial. Such experiments qualify as academic exercises. Yet, researchers with synthetic stimuli have proposed all numbers from 1 through 9 in different frequency bands as critical in pitch perception. Ironically, when the so-called critical band is removed from the sound, pitch remains the same. Watkins and Dyson[51(p74)] underscored the futility of the search because "any one musical note may be realized by a large number of sounds differing widely in their (Fourier) frequency composition although they all generate the same pitch percept." The body of experimental evidence in this study not only confirms the position but also provides additional insight thanks to invariance.

The mention of invariance ushers ecological psychologists on to the scene. Cutting[52(p202)] reported a major difference in the ways the "information-processing psychologist" and the "ecological psychologist" approach perception. For the former, the way an organism interacts with information must be approached with the aid of *computations* and representations; the latter, on the other hand, contends that the information must be picked up *without appeal to computations*. Computational approach to pitch perception has been unrewarding. Let us consider the position of ecological psychologists who reject computations in perception. An analogy will clarify the point.

Consider a case in vision. Walk into a musical instrument store and examine the vast array of different varieties of instruments. They do not mix in the eye however close one might be to another. The ear does not see what the eye sees in the store. However, if all the instruments are played together, the wave they create in the medium mix and produce a 'chaotic and meaningless acoustic jumble[18,53-55]' In contrast, the drum, the guitar, the xylophone bars, the trumpet, the human voice, and all the characteristics of each sound source, do not mix in the ear. If they did as in visual representations of sound, music would be altogether different. Thus, although the ear does not see the materials of sound sources in the store, the ear *sees* the sound sources when they are presented to the ear in the format that the ear can *see* and analyse—sound. Small[56(p46)] noted that "The auditory system, for the purposes of pitch perception, utilises certain aspects of the signal and neglects certain other aspects. If the signal could be described in the same terms as the auditory system "sees" it, then that description could be used to predict pitch perception for any type of signal." The present study has attempted to meet this requirement by establishing the isomorphism between a sound source and its sound—both being two forms of the same thing—the one for the eye, and the other for the ear, helping all organisms perceive the environment by eye and by ear (among other sensory means). These facts explain the way auditory perception mirrors visual perception.

Based on the above facts, this author sympathizes with Cutting [52(p214)] among many other ecological psychologists who assert that the ear does *not* do any computations but operates on invariants because "invariants unburden computational requirements on the perceiver[11,12,16-18,23-25,31,36,39,40,52,56-62]." More recently, Fowler & Hodges [63(p4)] contended against mental computations in perception and cognition because perception of the ecological niche is "direct" and "unmediated." The plausibility of the claims notwithstanding, they have no impact whatsoever on the information-processing scientist. After all, although scientific wars may be waged with words, none can be won by words but by experimental evidence. No ecological psychologist has tendered a mechanical, biological, acoustic, or neurophysiological evidence of perceptual invariance to bolster the claims and highlight fallacies and the futility of computational and information-processing approaches to perception without invariance. Therefore, in theory, ecological psychologists have abandoned the computational approach, acknowledge the importance of the sound source, and the need for invariance. However, the holistic view of the sound source as the stimulus, and its accessibility to the organism raise serious concerns that have led to the description of ecological acoustics as "qualitative, … have limited predictive power, based largely on analogy to vision[64(p3464)]." Other production-based experimenters focus how mechanical parameters individually impact vibrational frequency or the spectrum[65-70]. Such eclectic approaches, though well-intentioned, are caught in a vicious cycle because the spectrum does not portray mechanical traits that underly perception. Attempts to validate such findings have led researchers to adulterate invariance by introducing qualifiers such as "level-invariant", "more invariant" and "highly invariant[18(p381)]", "the reverse side of relativity," [71(p16)] "relative invariance,[72(p234)]", "contextual invariance" and "dynamic relative invariance[73(p182)]", "patterned stimulation[74(p19)]", etc. Besides, the term 'steady-state', which describes the absence of change in a spectral display, is often interpreted as invariant. Maeda[69] cautioned that invariance refers to the relationship between the object and perception as described at the outset of this study. The invariant does not call for a qualifier because the term *invariant* qualifies itself—it is either invariant or not, otherwise one square would be *squarer* than another in geometry. Controversies arising from the absence of a working principle are tearing ecological psychologists apart[61,62,75]. .

In contrast, the present study considers very closely the communication chain. A statement credited to Archytas (4th century A.D.) states that no sound can exist "without the striking of bodies against one another[1(p5)]." This fact underscores the primacy of the sound source in auditory perception as demonstrated in the present study through strict adherence to invariance. The established mechanical invariant accounts for perceptual constancy despite dramatic mechanical and acoustic variability. The experimental evidence in this study is sufficient not only for a theory of pitch/tone perception but also for a new foundation for hearing research. Here are some fundamental facts:

Figure 4(a) shows that the signal AT1 was produced without a balanced force (*F*ex) except what was necessary for the membrane to vibrate. Because the force cannot sustain vibration, the membrane collapses shortly after excitation. In contrast, the increased force in the reduced diameter of Drum C drives the membrane and sustains vibration; as the force dissipates, the membrane gradually returns to its position of rest. This applies also to strings. As the force in the string rises, whether through increased balanced force or reduced string length, the string loses its flexibility and the ability to create nodes and anti-nodes that generate Fourier components and harmonic spectra. Therefore, only sources with force around the middle of the force spectrum can produce regular waveforms and harmonic spectral structures. Signals generated by sources in regions of the two extremities of the force spectrum must be anharmonic[76]. Consequently, the same spectral composition resulting from different mechanical configurations of the sound source will elicit the perception of different percepts and *vice versa*. Without bleeding the string of the force arising from reduced string length, the full string and its sub-lengths in the experiment reported above would have produced higher pitches at Octave intervals. The researcher would strive in vain to explain the different pitches of the signals in terms of changing spectral structures. However, by displacing the extra force that determines higher pitches, it is clear that the associated spectral changes do not impact pitch since all the signals generate the same pitch. Many investigators endured the defects of psychoacoustics because there was no efficient alternative approach. Now that we know and have an alternative that is founded on invariance, comprising natural sources that produce real sounds that are used by real listeners in the real world, let us consider the implications of this new development on the future of hearing research.

Accepting that the force in a string is the stimulus to pitch raises overwhelming challenges. However, when vibrational frequency was thought to determine pitch, there was no technology for frequency extraction. Today, we have the technology. Similarly, if the experiments in the present study had been conducted 2,500 years ago, theories and practices in hearing research today would be different altogether. And that is what it could be from now on. To move forward in the understanding of brain functions in auditory perception, and provide real help for sufferers of hearing loss, we must learn to unlearn all the unprofitable things we have learned, and return to the sound source and learn to discover the properties that are transmitted to the organism as auditory codes in music and speech. The first major revision concerns the concept of force in a vibrating body. To facilitate the discussion, let us take a string.

In pre-historic physics of sound, the pull force exerted on a string is the force in the string. We know now that a string has its inherent force (*F*in)[77] besides the balanced force. Engineers employ different methods—stretching, compression, bending, etc —to assess the size of force a body can withstand before it yields[78,79]. Do these methods determine the force in the body itself? String equations leave the impression that string tension may be calculated from vibrational frequency, but it does not work[80,81]. Even if the pull force were the only source of force in the string, increasing frequency of sub-lengths of a string cannot be proportional to the square root of the tension that is held constant. Above all, no one, as far as this author knows, has ever demonstrated that a string actually vibrates at the calculated frequency and produces a pitch that corresponds to the frequency scale of musical notes. Sauveur[82], for example, coined the name *acoustics* for his purpose to transform music practice into a superior science. His pitch numbers were based on conjectures and assumptions. According to Lindsay[83(p63)], Sauveur “calculated by a somewhat dubious method the frequency of a given stretched string from the measured sag of the central point." These and a host of other discrepancies over the functional elements of hearing research portray it as physics and mathematics for academic exercises only[25,31,84,85].

An area of hearing research that has witnessed “the most dramatic change is in the field of tools rather than in the understanding of the "speech code."”[86(p788)] In fact, the progress has been described as a deceptive “boom.”[87(ppi,ii)]. In contrast, the insights from this study demand accurate evaluation of force in a string toward a viable psychological pitch scale[51,88-89]. The success would offer phenomenal insights on how the auditory mechanism processes sound to extract and measure pitch. The focus on the sound source would reduce or even cease investing more material and human resources in the search for pitch in the spectrum where it does not exist[25,39,61,75,90]. Hearing is psycho-mechanical; the primary object is the sound source[25].

4. Conclusion

The Pythagorean scientific study of hearing was a perfect but incomplete experiment in the absence of an invariant mechanical determinant of pitch. The deviation into acoustics has proved to be an unprofitable distraction. The present study has introduced an invariant—the hitherto missing psychological element—into hearing research to transform it from a purely physical and mathematical science into a science of auditory behaviour. The experimental data support the following conclusions:

(1) Pitch

(a) Pitch is the degree of force in a sound source.

(b) A sound source (string or membrane) calls for balanced force (*F*ex) only if the force that is the inherent property of the source (*F*in) is insufficient for the desired pitch;

(c) The force in a string or membrane is not constant when its dimensions are modified even though the balanced force is held constant;

(d) Contrary to pre-historic physics of sound, the force in a body is adjustable through modifications to its physical dimensions, or via the balanced force exerted on the body.

(e) The current foundation of hearing research is unsuitable for a science of auditory behaviour;

(f) Unlike acoustic wave, sound source properties do not mix in the ear; therefore the ear does not perform any computations in pitch/tone perception;

(g) Because the ear evaluates the force in the sound source to extract pitch and tone, perception is 'direct' and 'unmediated'.

(h) The psycho-mechanical evidence in this study proposes only one auditory mechanism in pitch and tone perception;

(i) The mechanical origin of pitch gives primacy to the sound source in hearing research;

(2) Phonemic tone

(a) A phonemic tone in Yoruba comprises a cluster of pitches that derive their autonomous tonal quality from the degree of force in the sound source;

(b) Tones are perceived on an absolute judgment basis in any pitch environment regardless of the vibrational frequency of the sound source which is the joint product of physical, mechanical, and other properties of the source;

(c) Pitches within tone boundaries are tonally ambiguous; they embody the qualities of adjacent tone categories.

(d) The multi-pitch structure of a tone ensures that pitch changes by several steps may not provoke the perception of a different tone. Therefore, tone language users can exploit pitch changes to express emotions without compromising the lexical role of tone;

(3) Implications for Future Research

Mechanical, acoustic, and neurophysiological investigations into music and speech perception outside the framework established in this study on the platform of invariance will most likely always prove to be futile.

Conflicts of Interest Statement:

None

Acknowledgements Statement:

The drum experiments reported in this study are drawn from my doctoral thesis[26]. My deep gratitude goes to my supervisor Jacqueline Vaissière, Laboratory of Phonetics, Centre of Applied Linguistics and Phonetics, Sorbonne University (Paris III), and to Sinji Maeda, National Centre for Scientific Research (CNRS), Paris, for his counsels and technical support. I am also indebted to Joy Oluchi Uguru, president of the Acoustical Society of Nigeria, for encouraging my research through contact with researchers across the globe. I am thankful to L. Smith, Chair of the European Society of Medicine, for relentlessly encouraging me to prepare this paper for the Society's Special theme on neurophysiology. All errors in the presentation or interpretation of data remain my responsibility.

Funding Statement:

This research work was financed in full with funds provided by the family of Herr Robert & Frau Maria, Jockisch of Düsseldorf, Germany. We, at the Ecological Auditory Research Laboratory, London, UK, seize this opportunity to express sincere appreciation and admiration for their love for science. It is hoped that their sacrifices open the way to the long-awaited integration of the deaf and hard-of-hearing into the hearing world.

References:

1. Carterette, E. C. Some historical notes on research in hearing, in Carterette, E. C., Friedman, M. P. (eds.) *Handbook of Perception*. New York: Academic Press: 1978.

2. Stephen, R. W. B., Bates, A. E. *Wave Motion and Sound*, London: Edward Arnold & Co: 1950.

3. Ohm, G. S. Ueber die Definition des Tones, nebst daran geknupfter Theorie der Sirene und ahnlicher tonbildenden Vorrichtungen. *Poggendors Annalen der Physik und Chemie*: 1843;59:513–565.

4. Helmholtz, H. L .F. *On the Sensations of Tone as a Physiological Basis for the Theory of Music,* Translated from the German Version of 1877 and Revised by Ellis, A.J. 1954), New York: Dover Publications: 1877.

5. de Cheveigné, A. Pitch Perception Models, in Plack, C. J., Oxenham, A. J., Fay, R. R. (eds.) *Pitch: Neural Coding and Perception. Springer Handbook of Auditory Research.* London: Springer: 2005:169-233.

6. Winter, I. M. The neurophysiology of pitch, in Plack, C. J., Oxenham, A. J. (eds.) *Pitch: Neural Coding and Perception, Springer Handbook of Auditory Research.* New York, Springer: 2005.99-146.

7. Moore, B. C. J. Aspects of auditory processing related to speech perception, in Hardcastle, W. J., Laver, J., Gibbon, F. E. (eds.) *The Handbook of Phonetic Sciences,* Sussex: Wiley-Blackwell: 2010:454-488.

8. Békésy, G. von. The missing fundamental and periodicity detection in hearing. *Journal of the Acoustical Society of America.* 1972;51:631-637.

9. Plack, C. J., Oxenham, A. J., Fay, R. R. (eds.) *Pitch: Neural Coding and Perception. Handbook of Auditory Research.* London: Springer: 2005.

10. Perkell, J. S., Klatt, D. H. (eds.) *Invariance and Variability in Speech Processes,* New Jersey: Lawrence Erlbaum Associates: 1986.

11. Essien, A.J. Invariance: Its Nature and Role in Auditory Psychophysics. In Essien, A. J. (author) *Sound Sources*, Michael Terence Publishing: London: 2019:35-68.

12. Stevens, S. Mathematics, Measurement, and Psychophysics, in Stevens, S. S. (ed.) *Handbook of Experimental Psychology.* London: John Wiley & Sons: 1960:1-49

13. Bell, E. T. The Development of Mathematics. London: MacGraw-Hill: 1945.

14. Gibson, J. J. *The Ecological Approach to Visual Perception.* Boston. MA: Houghton-Mifflin: 1979.

15. Boring, E. G. The stimulus error. *The American Journal of Psychology*: 1921; 32:449-471.

16. Essien, A. J. Pitch in quantitative psychology. In Essien, A. J. (author) *Sound Sources*, Michael Terence Publishing: London: 2019:99-141.

17. Fechner, G. *Elements of Psychophysics Vol.* I. Translated from the German by Adler, H.E. (l966), London: Holt, Rinehart and Winston: 1860.

18. Cariani, P., Micheyl, C. Toward a Theory of Information Processing in Auditory Cortex, in Poeppel, D., Overath, J., Popper, A. N., Fay, R. R. (eds.) *The Human Auditory Cortex. Springer Handbook of Auditory Research,* London: Dordrecht: 2012:351-390

19. Stevens, S. S., Davis, H. *Hearing: Its Psychology and Physiology*, London: John Wiley & Sons: 1938.

20. Wilson, C. E. *Noise Control (Revised Edition).* Malabar, Florida: Krieger Publishing Company. 2006

21. Berg, R. E., Stork, D. G. *The Physics of Sound*. New Jersey: Prentice Hall: 1982.

22. Akrill, T. B., Miller, C. J. *Mechanics, Vibrations and Waves*, London: J. Murray. 1974.

23. Öhman, S. E. G. What is it we perceive when we perceive speech? in Cohen, A. and Nooteboom, S. G. (eds.) *Structure and Process in Speech Perception*, Berlin: Springer-Verlag. 1975:36-47.

24. Essien, A. J. Ecological Conception of Sound. In Essien, A.J. (author) *Sound Sources: The Origin of Auditory Sensations*. Michael Terence Publishing. London: 2019:191-221.

25. Essien, A.J. *Sound Sources: The Origin of Auditory Sensations*. Michael Terence Publishing. London: 2019.

26. Essien, A. J. *The Perception of Yoruba Tone: Experimental Evidence from Drums, Speech Signals and Synthesis* (In French). Unpublished Ph.D. thesis. Institute of Phonetics, Sorbonne University, Paris III. France: 2000.

27. Essien, A. J Pitch, tone, and intonation: The co-existence of tone and intonation in a tone language, Yoruba. Work in Progress: 2023

28. Akpabot, S. Standard drum patterns in Nigeria. *African Music*: 1971;5(1):37-39.

29. Newton, I. *Philosophiæ Naturalis Principia Mathematica*, (*Mathematical Principles of Natural Philosophy*), Vols. 1 and 2 Translated into English by Andrew Motte (1729). Revised and supplied with a historical and explanatory appendix by Florian Cajori. New York: Greenwood Press: 1687.

30. Radocy, R. E. Some unanswered questions in musical perception, *Contributions to Music Education:* 1978;(6):73-81.

31. Essien, A. J. The mechanical invariance factor in musical acoustics and perception (revisited). *American Journal of Modern Physics,* 2018:(7):1-13.

32. Békésy, G. von. Hearing theories and complex sounds. *Journal of the Acoustical Society of America*: 1963;(35):588-601.

33. Metzger, W. Consciousness, perception and action, In Carterette, E. C. and Friedman, M. P. (eds.) *Handbook of Perception,* New York: Academic Press: 1974;(1):109-122

34. Neisser, U. *Cognition and Reality*, San Francisco: Freeman. 1976.

35. Diehl, R. L. Coproduction and direct perception of phonetic segments: A critique, *Journal of Phonetics*: 1986;(14):61-66.

36. Port, R. F. Invariance in phonetics, in Perkell, J. S. and Klatt, D. H. (eds.) *Invariance and Variability in Speech Processes,.* New Jersey: Lawrence Erlbaum Associates: 1986:540-558

37. Householder, F. W. Jnr Review: Visible Speech by Potter, R.K., Kopp, G.A., Green, H.C. (1947): *Word:* 1948;(4):53-57.

38. Bregman, A. S. *Auditory Scene Analysis: The Perceptual Organization of Sound,* Cambridge, Massachusetts: The M.I.T. Press: 1990

39. Fischer-Jørgensen, E. What can the new techniques of acoustic phonetics contribute to linguistics? in Saporta, S. (ed.) *Psycholinguistics: A Book of Readings*, New York: Holt, Rinehart and Winston: 1961:112-142.

40. Mol, H. *Fundamental of Phonetics*. The

Hague: Mouton & Co.: 1963.

41. Mol, H. *Fundamental of Phonetics, Part II.* The Hague. Mouton & Co: 1970.

42. Fant, G. *Speech Sounds and Features*, Cambridge, Massachusetts: The M.I.T. Press: 1973.

43. Boomsliter, P. C., Creel, W. Research potentials in auditory characteristics of violin tone. *Journal of the Acoustical Society of America*: 1972;(51):1984-1993.

44. Fletcher, H. *Speech and Hearing*, London: Macmillan: 1929.

45. Plomp, R. Pitch of complex tones, *Journal of the Acoustical Society of America*: 1967;(41):1526-1533.

46. Patterson, R. D., Wightman, F. L. Residue pitch as a function of component spacing, *Journal of the Acoustical Society of America*: 1976;(59):1450-1459.

47. Terhardt, E., Stoll, G., Seewann, M. Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions, *Journal of the Acoustical Society of America*: 1982;(71):671-678.

48. Houstma, A. J. M. and Smurzynski, J. Pitch identification and discrimination for complex tones with many harmonics, *Journal of the Acoustical Society of America:* 1990;(117):1326-1336.

49. Terhardt, E. Information-bearing elements of complex sounds, *Journal of the Acoustical Society of America:* 1992;(91):2357.

50. Hartman, W. M., Doty, S. L. On the pitches of the components of a complex tone, *Journal of the Acoustical Society of America:* 1996;(99):567-578.

51. Watkins, A. J., Dyson, M. C. On the perception organization of tone sequence and melodies, in Howell, P., Cross, I. and West, R. (eds.) *Musical Structure and Perception*, New York, Academic Press: 1985:71-119.

52. Cutting, J. E. Two ecological perspectives: Gibson vs. Shaw and Turvey, *The Am. J. Psychology*: 1982;95(2):199-222

53. Nooteboom, S. G. Chairmans opening speech: Seminar on the Perception of Temporal and Intonational Patterns of Speech, in Fant, G. and Tatham, W.A.A. (eds.) *Auditory Analysis and Perception of Speech*, New York: Academic Press: 1975

54. Bloothooft, G., Hazan, V., Huber, D. and Llisterri, J. Introduction, in Bloothooft, G., Hazan, V., Huber, D. and Llisterri, J. (eds.) *European Studies in Phonetics and Speech Communication*: Utrecht: OTS Publications: 1995:xiii-xv.

55. Hirsh, I. J., Watson, C. S. Auditory psychophysics and perception, Annual Reviews of Psychology: 1996;(47):461-484.

56. Small, A. M. Periodicity pitch, in Tobias, J. V. (ed.) *Foundations of Modern Auditory Theory*. 1970;(1):3-54.

57. Mace, W. M. James J. Gibson's Ecological Approach: Perceiving What Exists, *Ethics and the Environment:* 2005;10(2):195-216.

58. Fowler, C. A. Sound-producing sources as objects of perception: Rate normalization and nonspeech perception, *Journal of the Acoustical Society of America:* 1990;(88): 1236-1249.

59. Repp, B. H. Categorical perception: Issues, methods and findings, *Haskins Laboratories Status Report on Speech Research*, 1982;(70):99-183.

60. Neuhoff, J. G. Preface, in Neuhoff, J. G.

(ed.) *Ecological Psycho-acoustics.* London: Elsevier: 2004.

61. Van Valkenburg, D. and Kubovy, M. From Gibsons fire to Gestalts: A bridge-building theory of perceptual objecthood, in Neuhoff, J. G. (ed.) *Ecological Psychoacoustics,* London: Elsevier: 2004:113-147

62. Whalen, D. H. Three lines of evidence for direct links between production and perception in speech, *Proceedings l4th International Congress of Phonetic Sciences*: San Francisco: 1999:1257-1260.

63. Fowler, C. A., Hodges, B. Finding common ground: Alternatives to code models for language use, *New Ideas in Psychology.* 2016.doi.org/10.1016/j.newideapsych.

64. Yost, W. A. Perceiving sounds in the real world: An introduction to human complex sound perception, *Frontiers in Bioscience,* 2007;(12):3461-3467

65. Obata, J. and Tesima, T. Experimental studies on the sound and vibration of drum, *Journal of the Acoustical Society of America.,* 1935;(6):267-274.

66. Lewis, D. & Cowan, M. The influence of intensity on the pitch of violin and "cello" tones, *Journal of the Acoustical Society of America:* 1936;(8):20-22.

67. Titze, I. R. Vocal fold mass is not a useful quantity for describing fo in vocalization, *Journal of Speech, Language, and Hearing Research,* 2011;(54):520-522.

68. Gaver, W. W. How do we hear in the world? Exploration in ecological acoustics, *Ecological Psychology*: 1993;(5):285-313.

69. McNeil, L. E., Mitran, S. Vibrational frequencies and tuning of the African mbira, *Journal of the Acoustical Society of America:* 2008;(123):1169-1168.

70. Maeda, S. Links between speech production and acoustics: A sketch, Proceedings 12th Inter. Cong of Phonetic Sciences, Aix-En-Provence, France: 1991;(1): 200-205.

71. Jakobson, R. and Waugh, L. *The Sound Shape of Language*, Brighton: Harvester Press: 1979.

72. Kent, R. D. The iceberg hypothesis: the temporal assembly of speech movements, in Perkell, J. S. and Klatt, D. H. (eds.) *Invariance and Variability in Speech Processes,* New Jersey: Lawrence Erlbaum Associates: 1986:234-242.

73. Blumstein, S. E. On acoustic invariance in speech, in Perkell, J. S., Klatt, D. H. (eds.) *Invariance and Variability in Speech Processes*, New Jersey: Lawrence Erlbaum Associates: 1986:178-201.

74. Lutfi, R. A. Human sound source identification, in Yost, W. A., Popper, A. N. and Fay, R. R. (eds.) *Auditory Perception of Sound Sources. Springer Handbook of Auditory Research,* New York: Springer: 2008; (29):13-42.

75. Neuhoff, J. G. Ecological Psychoacoustics: Introduction and History, in Neuhoff, J. G. (ed.) *Ecological Psychoacoustics*, London: Elsevier: 2004:1-13.

76. Essien, A.J. Mechanics of spectral change. In Essien, A. J. (author) *Sound Sources*,. Michael Terence Publishing. London: 2019:385-446

77. Essien, A. J. The force/tension dichotomy. *In* Essien, A.J. (author) *Sound Sources*, Michael Terence Publishing. London: 2019:337-361.

78. Gillam, E. *Materials under Test*. London: Newnes-Butterworths: 1969.

79. Higgins, R. A. *Materials for Engineers and Technicians (5th Edition),* Amsterdam: Newnes-Elsevier: 2010.

80. Cross, R., Lindsey, C. and Andruczy, D. Laboratory testing of tennis strings, *Journal of Sports Engineering*: 2000;(3):219-230.

81. Cross, R. and Bower, R. Measurements of string tension in a tennis racket, *J. Sports Engineering*: 2001:4(3):165-175.

82. Sauveur, J. *Principes dAcoustique et de Musique, ou Système Général des Intervalles des Sons, et de son Application a tous les Instruments de Musique* (Paris). Reprinted l973*,* Geneva: Minkoff, 1701.

83. Lindsay, R. B. The story of acoustics, *Journal of the Acoustical Society of America:* 1966;(39)629-644.

84. Dowling, W. J., Harwood, D. L. *Music Cognition*. London: Academic Press: 1986.

85. Plack, C. J., Oxenham, A. J. Overview: The present and future of pitch, in Plack, C. J. and Oxenham, A. J. (eds.) *Pitch: Neural Coding and Perception. Handbook of Auditory Research,* New York: Springer, 2005:1-6.

86. Carlson, R., Granström, B. Speech Synthesis, in Hardcastle, W. J., Laver, J. (eds.) *The Handbook of Phonetic Sciences*, Oxford: Blackwell Publishers: 1997:768-788.

87. Perkell, J. S., Klatt, D. H. (eds.) *Invariance and Variability in Speech Processes,* New Jersey: Lawrence Erlbaum Associates: 1986.

88. Stevens, S. S., Volkman, J. The relationship of pitch to frequency: A revised scale, *The American Journal of Psychology:* 1940;(53):329-353.

89. Siegle, R. J. A replication of the mel scale of pitch, *American Journal of Psychology*: 1965;(78):615-620.

90. Read, J. C. A. The place of human psychophysics in modern neuroscience. *Neuroscience*, 2015;(296):116-129.