Music Perception; Paper ICA2016-1179

The mechanical invariance factor in musical acoustics and perception

Akpan J. Essien

Acoustical Society of Nigeria, University of Nigeria, Nsukka, Nigeria

Abstract

Acoustical and neurophysiological investigations into pitch perception repose on the Pythagorean string ratio theory of pitch interval. The validity of the theory has been denied recently on the platform of Invariance. Essien (2014) demonstrated experimentally that, contrary to established tradition in physics of sound, string tension is not constant but varies inversely with string length even though the oppositely-directed force exerted on the string is held constant. The finding called for complete review of theories and practices aimed to unravel the principle of the auditory mechanism. The present paper reports on the impact of a string’s force of resistance to deformation on the string’s vibrational frequency, spectral structure and change. Sub-lengths of a string are shown to have very little or no effect at all on a string's vibrational frequency and pitch. The data exposed to account refute the string ratio theory of pitch interval; they portray the force in a string as the mechanical parameter in control of spectral structure and pitch. Implications for future research in musical acoustics and perception are discussed.

Keywords: Invariance; Pitch Perception, Mechanics of spectral change.
The mechanical invariance factor in musical acoustics and perception

1. Introduction

Acoustic cues for perception “represent a narrow-minded way of thinking which leads us into a blind alley when faced with the problem of discovering the real nature of the auditory mechanism.” [1]

The Pythagorean string ratio theory is the pivot around which revolve all work in hearing [2; 3]. Ohm’s Acoustical Law [4], and Helmholtz’s Resonance/Place Theory [5], both reflecting perception by ratios of stimulus magnitude, are the acoustical and neurophysiological facets of the string ratio theory. The difficulties encountered in the search for acoustic cues to perception, whether in music or speech, have led many investigators to label the objective “elusive” [6] or “ambitious” [7]. The ecological approach is presumably promising [8, 9] However, modern investigations seek after the impact of individual parameters of the sound source on the physical parameter frequency of vibration, and not the invariant property of the sound source that underlies subjective pitch. [10, 11, 12, 13]. In this regard, Essien [14] pointed out the existence of a force that is the inherent property of a string (hereafter Fin) outside the oppositely-directed force that is applied externally to a string (hereafter Fex). The former was shown to vary inversely with string length, such that the force in the string, in reality, is not constant but varies with pitch even though the latter (called tension) is held constant. This paper focuses the impact of that force on a string’s vibrational frequency, spectral structure and subjective pitch. Implications for future research in pitch production, musical acoustics, and perception are discussed.

2. Experiment

2.1. Equipment and Method

The equipment used in this study was a sonometer; it was described in detail in Essien [14]. Only the 85-gauge string was involved in the present experiment. The size of the oppositely-directed force exerted on the string was measured by means of an integrated spring balance. Effective string length was adjusted by means of a mobile bridge. The resonator was equipped with a built-in pickup socket for a sound-recording device. The informant was a female absolute pitch professional classical violinist who is identified in this work as YI. She was taught music from infancy, and had practised the art for 23 years at the time of the experiment. She held a Master's degree from the Royal Academy of Music, London, and was teacher of music (classical violin) at Pimlico Academy, London, UK. The task was as follows:

Given the full string as 860 mm long, the informant was required to halve it successively; and to tune each sub-length to the same pitch, i.e. A3. The operation produced six string lengths—the full string (860 mm) and five sub-lengths measuring 430, 215, 107, and 53 mm.
As the above task shows, the experiment involves sub-lengths of the same string. Nevertheless, it is readily noticeable that it does not address pitch interval (which implies perception of two different pitches) but focuses the perception of pitch per se. Because the full string and its sub-lengths are tuned to the same subjective pitch, the experiment operates on the platform of invariance, seeking to detect, from among the different mechanical configurations of the strings, the parameter that remains inseparably tied to pitch [15, 16]. The six signals produced at each of the six string configurations were recorded on tape using an AIWA hi-fi cassette deck.

2.2. Acoustic analysis
The signal analysing software Sigview was used for the spectral analyses. It is necessary to go through some samples for better acquaintance with terminology and clearer understanding of the summary of results. Consider, for example, the two signals in figure 1. To highlight the point of this paper, the signal (bottom) is a portion extracted from the signal which produced by the full string (860 mm long); the complete signal is 12.2 secs long. The signal (top), produced by a sub-length of the same string measuring 53 mm, is 0.3 secs long. Nevertheless, these signals generate the same subjective musical pitch A_3. Let us consider the spectral characteristics of the two signals.

The vibrational frequency of the signal at bottom of picture as obtained by manual peak-to-peak estimation was 55.55 Hz. The result of FFT spectrum analysis of the same signal is presented in figure 2 (top). The 165 Hz. component with the highest intensity at 45.75 dB is considered as the resonant frequency of the signal, but it is not the fundamental because the other components are not integral multiples of the 165 Hz. partial. Therefore, if the harmonic ruler is placed at the 165 Hz. component, a harmonic marker falls on every third component of the spectrum, thus rendering the entire structure non-harmonic. If, however, the ruler is set further to the left as shown in the top picture, all the partials are in perfect harmonic arrangement. The component that makes this harmonic arrangement possible is the 55 Hz. component labelled 'Missing Fo'. Therefore, the fundamental for this sound is not the imposing 165 Hz partial but the 55 Hz. component which, ironically, is missing from the spectral structure of the sound.

![Figure 1: The impact of mechanical properties of a string on the waveforms of naturally produced stimuli. The two signals were produced by different mechanical configurations of a string; they generated the same subjective musical pitch A_3.](image-url)
Figure 2: Physical manifestations of an A₃ musical tone produced on two different string configurations. Top picture is the spectrum of the signal by the 860 mm-long string. Middle picture is the spectrum of the 53 mm-long string. The bottom picture presents the FFT analysis of the 53 mm-long string. The two signals produce the same pitch.
The middle picture is the spectrum of the signal produced by the 53 mm-long string. In contrast, this signal manifests the absence of periodic vibrations and thus defies a fundamental acoustic description of musical sounds. However, the highlight is the dominant composite component with two peaks, one at 220.72 Hz. and the other at 228.29 Hz. The picture (bottom) presents the normalised FFT analysis of this signal; it points to the existence of a resonant frequency component at 226.1 Hz. Thus, different acoustic analyses can lead to altogether different results. Let us determine the contribution of the spectral data to the unchanging musical pitch of the two stimuli through an examination of their constancy (or otherwise) in the course of each signal’s evolution.

2.3. Evolution of spectral components and structure

Naturally produced sounds are characterised by change as a function of time; this instability poses a defiant problem for psychoacoustic theories of music and speech perception [17, 18, 19]). In this examination, we shall view each signal through spectral slices for constancy and variability in the behaviours of the dominant spectral components of interest. All the slices were adjusted to fit on a common amplitude scale for better observation of their relative intensity. Figure 3(a) presents six spectral slices of the A3 musical tone produced by the 860 mm-long string. The 220 Hz. partial appears in the third slice as a subsidiary peak. In the fourth slice, this component is in strong competition with the 164.19 Hz. partial. It dominates over the 165.54 Hz partial, but is dominated over by the 164.86 Hz. component in slice 6. Figure 3(b) presents four slices showing the spectral behaviour of the A3 tone produced on the 53 mm-long string. The four spectral slices show no trace of the 220 Hz. component in the middle picture of figure 2.

Spectral components and structures in 2D presentations offer only a partial view of the reality. For a better appraisal of the perceptual import of the components, figure 3c presents the spectral structures of the two signals under study in 3D format. The spectrogram (left) is the evolution of slice 4 in figure 3(a). The imposing spectral component is the 165 Hz. partial. This component shows itself as the dominant component of the signal in terms of intensity and duration. Strangely, the 220 Hz. partial next to it does not manifest the competitive height that was recorded in the 2D presentation in slice 4 of figure 3(a) even in the region of high energy; there is not much of this component thereafter. As regards the spectrogram of the signal produced on the 53 mm-long string (bottom right), we note the first component which runs from start to end of the signal. The 220 Hz. frequency component was not manifest in any of the four spectral slices of the signal. The 226 Hz. frequency component belongs to the second formant which is shown in the spectrogram (bottom right) as short-lived; it decays and disappears within the first quarter of the signal. It has to be recalled that the two peaks were branches on one stem. Thus, the frequency of 226 Hz is only one of the frequency components in this formant. It was picked up only at the end of the signal as shown in figure 3(b).

In the two signals under examination, the presence of the 220 Hz component in each signal is of interest. But because the component is neither the fundamental nor the resonant frequency of the signal makes it difficult to attribute to it the pitch of the signal in accordance with Ohm’s acoustical law [4] or Helmholtz’s resonance/place theory [5]. In the case of the harmonic signal in figure 2a, the 220 Hz. partial is the fourth component; and besides, it has very little energy and is short-lived. Why would the ear not choose the 165 Hz partial which we would undoubtedly produce maximal stimulation on the basilar membrane? The theoretical possibility to retrieve the 55 Hz. fundamental in this circumstance poses even a more difficult problem since the signal is defined in auditory terms as an A3 musical tone; specified acoustically by the frequency of 220 Hz. The recorded fundamental frequency of 55 Hz. should produce a pitch two octaves lower i.e. A1 rather than A3. The controversies arising from this phenomenon (called the missing fundamental, or residue pitch, or periodicity pitch, or low pitch) permeate every fibre of hearing research and raise questions that have never been answered, neither at the mechanical level [11], nor at the acoustic level [20, 21, 22], nor at the neurophysiological level [23, 24, 25, 26]. Because the fundamental is missing, the perception is not direct but is purportedly derived through computations [27].
(a) Spectral slices of an A₃ tone from the 860 mm-long string

(b) Spectral slices of an A₃ tone from the 53 mm-long string

(c) 3D spectral images of A₃ tones by the 860 mm string (bottom left) and the 53 mm string (bottom right)

Figure 3: Spectral slices showing the behaviours of spectral components of two signals of the same pitch: (a) the A220 tone produced on the 860 mm-long string; (b) The A220 tone, produced on the 53 mm-long string. partials of a signal in 2D presentation. The signal is an A220 Hz. musical tone produced on a string.
However, it has been argued, and plausibly too, that the ear does not seem to function in that manner \(1, 3, 20, 21, 22, 23, 24\). We cannot dwell on disputes regarding the unbridged chasm between auditory experience and physical characteristics of stimuli. In the present case, since the 220 Hz component is present in the two signals examined above, proponents of psychoacoustic theories of hearing might find in it a flimsy string on which to hang the weighty hope to explain music, speech and hearing by purely psychoacoustic procedures. To take a stand on these issues, let us turn attention to the results of the string tuning experiment described above. The results will be examined from two different standpoints: (1) From the viewpoint of psychoacoustics; (2) From the viewpoint of invariance.

2.4. Results

2.4.1. The Invariance Problem in Psychoacoustics

Table 1(a) summarizes the results of the experiment on the basis of the Pythagorean string ratio theory. String length was halved successively four times. The size of the oppositely-directed force \((F_{ex})\) needed for the 860 mm-long string to sound the musical tone \(A_3\) was 22 kg. This size of force was held constant in accord with the terms of the string ratio theory. Theoretically, each sub-length should produce a pitch an octave higher—\(A_4, A_5, \text{ etc.} \) (Column 3). The term Resonant Frequency (column 4) refers to the spectral component with the highest intensity, and Partial (column 5) specifies the position of the harmonic component in the spectrum. The \(F_o\) (column 6) denotes the lowest component of which all the other components in the spectrum are integral multiples as described earlier. Thus, entries in Row 1 state that the 860 mm-long-string, tensioned to 22 kg, has the most prominent partial at 165 Hz, and this component being the 3rd partial in the signal. The \(f_o\) for this sound, as established earlier above, is 55 Hz. In this way, \(f_o\) values were established for all signals (where possible). The data show irregular rising resonant frequency values as string length decreases. This is a well-known phenomenon. However, Column 6 shows no doubling of \(f_o\) when

| Data Row | String (mm) | \(F_{ex}\) (kg) | Musical Pitch | Res. Freq. (Hz) | Partial |
|----------|-------------|----------------|---------------|----------------|---------|
| 1        | 860         | 22             | \(A_3\) 165   | 3rd            | 55      |
| 2        | 430         | 22             | \(A_3\) 278   | 3rd            | 93      |
| 3        | 215         | 22             | \(A_3\) 373   | 2nd            | 185     |
| 4        | 107         | 22             | \(A_3\) 372   | 1st            | 372     |
| 5        | 53          | 22             | \(A_3\) 845   | 2nd            | N/A     |

Table 1: String ratio approach vs Mechanical invariance approach to hearing. In (a), the Pythagorean string ratio approach to hearing attributes pitch increments to sub-lengths of a string. In (b) the mechanical invariance approach presents the force in the string as the parameter in pitch and frequency control regardless of string length.
string length was reduced to 430 mm. Thereafter, however, the \( f_0 \) is shown to double with the halving of string length. The precision is astounding! We note also the tendency on the part of the resonant component to shift backward in the spectrum as length of string decreases. The signal produced by the 53 mm long string had two competing resonant components, one at 845 Hz. and the other at 1566 Hz; the \( f_0 \) could not be established using frequency interval.

One might be tempted to draw precipitated conclusions on these results and attribute the observed spectral behaviours of the signals to length of string according to the string ratio theory. Within the framework of invariance, however, everything that we have examined above is devoid of psychological essence because they are all based on variants—different string lengths, different resonant frequencies, different \( f_0 \) values, different pitches, etc. From the viewpoint of invariance, all sub-lengths of the string must produce the same pitch. In the midst of mechanical variability, a parameter that maintains an unchanging relationship with the sensation pitch may be detected. Only the parameter that meets that criterion constitutes a psychological basis for a theory of auditory perception. Thus, the observed acoustic precision in table 1(a) is, sadly, at the root of all the disputes and controversies and, indeed, the failure of psychoacoustics in music, speech, and hearing research if string length is not in control of frequency and pitch. Let us consider the results of the mechanical invariance approach in table 1(b).

2.4.2. The mechanical invariance factor

Table 1(b) introduces invariance into the study. Therefore, all the different string lengths were tuned to the same musical pitch \( A_3 \). The new addition in table 1(b) which was described in the Introduction is the parameter \( F_{\text{in}} \) in column 4. Entries in row 1 report that the 860 mm-long string, tensioned to 22 kg, produced the musical pitch \( A_3 \) (or A220) at the resonant frequency of 165 Hz. The \( f_0 \) of the signal as established earlier is 55 Hz. We do not know the inherent force of the string (\( F_{\text{in}} \)) at this point. Then, string length was shortened to 645 mm while \( F_{\text{ex}} \) was held constant. The subjective pitch of signal rose as evidenced by the rise in frequency to 211 Hz. Because our focus is the mechanical determinant of the \( A_3 \) musical tone, the new and higher pitch and frequency are of no interest to the present experiment; therefore, it is labelled UP (Unlabelled Pitch). In row 3 the 645 mm-long string was tuned to the same \( A_3 \) musical pitch. To achieve that goal, \( F_{\text{ex}} \) was reduced to 13.2 kg. The size of \( F_{\text{in}} \) is measured in terms of the force displaced, i.e. 8.8 kg, for the relatively shorter string to sound the same pitch \( A_3 \). The \( f_0 \) was established as 55.7 Hz.

A close scrutiny of the table in this manner shows increased pitch and \( f_0 \) each time the string is shortened while \( F_{\text{ex}} \) is held constant. However, whenever a sub-length of string was re-tuned to the same musical pitch \( A_3 \), \( F_{\text{ex}} \) fell; and the \( f_0 \) (where applicable) returned to the same level, lying between 54.8 and 57 Hz. Interestingly, the shortest 53 mm sub-length required no \( F_{\text{ex}} \) to produce the \( A_3 \) tone. Thus, the full string and its sub-lengths are potential generators of the same subjective pitch at about the same \( f_0 \).

The above facts are diametrically opposed to the aggregates in the foundation of hearing sciences as portrayed by table 1(a). On the basis of the evidence exposed to account in table 1(b), string sub-lengths do not control a string’s vibrational frequency or pitch. The inherent force of the string rises when string is shortened even though the force applied externally to the string is held constant. Efforts at finding pitch by current psychoacoustic procedures has been crowned with failure. Haggard [28] would describe it as “the search for the […] spectre in the spectrum.” Indeed, a complete review of theories and practices based on the string ratio theory, Ohm’s acoustical law, and Helmholtz’s resonant/place theory of music perception and hearing is compelling if auditory research is to attain the status of a behavioural science and progress even by little from where Pythagoras left off some 2,500 years ago. [29]
3. Conclusions
This paper presented experimental data that permit five critical conclusions: (1) The higher frequencies and pitches arising from sub-lengths of a string are not determined by string length but by the rise in the force which is the inherent property of the string (Fin) even though the force applied externally to the string (Fex) is held constant. (2) The enigmatic nature of the sensation pitch in music or speech is attributable to the founding of hearing sciences on a variant physical parameter of sound rather than the invariant property of the sound source that underlies the auditory sensation pitch. (3) The probable existence of an invariant mechanical parameter of sound sources that underlies unchanging perception of pitch despite acoustic variability is a potential case for direct perception in opposition to computational theories of auditory perception. (4) By implication, until the invariant property of the sound source that underlies a given auditory sensation is detected, isolated, and given acoustic representation, the search for functional codes in music and speech in the physical representations of sound is premature. (5) The quest for the principle of the auditory mechanism through exploitation of unrelated stimulus properties are, from the standpoint of invariance, subject to futility.

4. References
[1] Mol, H. Fundamentals of Phonetics. Mouton & Co. The Hague. 1963:
[2] Carterette, E.C. Some historical notes on research in hearing. Handbook of Perception, Vol. 4, 1978, pp 3-39.
[3] de Cheveigné, A. Pitch perception models. Springer Handbook of Auditory Research, Vol. 24, 2005: pp 169-233.
[4] Ohm, G.S. Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene. 1843. und ähnlicher tonbildender Vorrichtungen. Ann. Phys. Chem. Vol 59, pp 513-565.
[5] Helmholtz, H.L.F. On the Sensation of Tone; as a Physiological Basis for the Theory of Music. Translated from the German version of 1877 and revised by Ellis, A.J. Dover Publications New York. 1877.
[6] Dorman, M.F.; Wilson, B.S. The design and function of cochlear implants: Fusing medicine. Neural science, and engineering. American Scientist, Vol 92 (5), 2004, pp 436-445.
[7] Cariani, P.; Micheyl, C. Toward a theory of information processing in auditory cortex. Springer Handbook of Auditory Research, Vol. 43, 2012, pp 351-390.
[8] Neuhoff, J.G.: Ecological Psychoacoustics: Introduction and History. In Neuhoff, J.G. Ed.: Ecological Psychoacoustics, 2004, pp 1-13.
[9] Yost, W.A.: Perceiving sounds in the real world. An introduction to human complex sound perception. Frontiers in Bioscience, Vol 12, 2007, pp 3461–3467
[10] Obata, J.; Tesima, T. Experimental studies on the sound and vibration of drum. J. Acoust Soc. Am. Vol 6, 1935: pp 267-274.
[11] Gaver, W.W. What in the world do we hear? An ecological approach to auditory event perception. Ecological Psychology, Vol 5, 1993 pp 1-29.
[12] Gaver, W.W. How do we hear in the world? Exploration in ecological acoustics. Ecological Psychology, Vol 5, 1993: pp 285-313.
[13] Carello, C.; Anderson, K.L.; Kunkler-Peck, A.J. Perception of object length by sound. Psychological Science, Vol 9, 1998, pp 211-214.
[14] Essien, A J. The tension theory of pitch production and perception. Proceedings of Second International Conference on Acoustics. Acoustical Society of Nigeria. University of Nigeria, Nsukka, Nigeria. 13-16 October 2014. (In Press).
[15] Fechner, G. Elements of Psychophysics. Vol. 1. Translated from the German by Adler, H.E. Holt, Rinehart & Winston. New York. London. 1860:
[16] Bell, E.T. The Development of Mathematics. MacGraw-Hill. New York. London. 1945.
[17] d’Alessandro, C., Rosset, S.; Rossi, J-P. The pitch of short duration fundamental frequency glissandos. *J. Acoust. Soc. Am.*, Vol 104, 1998, pp 2339-2348.
[18] Collier, R. Intonation analysis: The perception of speech melody in relation to acoustics and production. *European Conference on Speech Communication and Technology*. Vol. 1, 1989: pp 38-44.
[19] Moore, B.C.J. Aspects of auditory processing related to speech perception. *The Handbook of Phonetis Sciences*, 2010, pp 454-488.
[20] Radocy, R.E. Some unanswered questions in musical perception. *Contributions to Music Education*. Vol. 6, 1978, pp 73-81.
[21] Békésy, G. von. Hearing theories and complex sounds. *J. Acoust Soc. Am*, Vol 35, 1963, pp 588-601.
[22] Békésy, G. von. The missing fundamental and periodicity detection in hearing. *J. Acoust Soc. Am* 1972, Vol 51, pp 631-637.
[23] Small, A.M. Periodicity pitch. *Foundations of Modern Auditory Theory*, Vol. 1, 1970: pp 3-54.
[24] Winter, I.M. The neurophysiology of pitch. *Springer Handbook of Auditory Research*, Vol. 24, 2005, pp 99-146.
[25] Plack, C.J.; Oxenham, A.J. Overview: The present and future of pitch. *Springer Handbook of Auditory Research*, Vol. 24, 2005, pp 1-6.
[26] Plack, C.J.; Oxenham, A.J. The psychophysics of pitch. *Springer Handbook of Auditory Research*, Vol. 24, 2005; pp 7-55.
[27] Ullman, S. Against direct perception. *The Behavioural and Brain Sciences*, Vol 3(3), 1980, pp 373-415.
[28] Haggard, M.P. Understanding speech understanding. *Structure and Process in Speech Perception*, 1975, pp 3-15.
[29] Essien, A.J. *The Ecological Foundation of Hearing Sciences: The basis for theories of music, speech and auditory analysis*. New Generation Publishing Company, London. (In Press) 2016:

5. Acknowledgments
My sincere appreciation goes to J. Uguru, President of the Acoustical Society of Nigeria, University of Nigeria, Nsukka, for her encouragement to me to prepare this paper for the Congress. I also thank the Congress officials for helping me through the registration process. The paper was financed fully with funds from my family. Errors of judgment remain my responsibility.